目 录

第　一　章

关于算术级数的范德瓦尔登定理

§1

1928年夏天，我在哥庭根度过了几个星期。象往常一样，许多获得奖学金的外国学生到这儿来学习。他们中有许多人我早已认识，有一些还成了我的朋友。我到那里时，那儿的数学家们讨论的主题是年轻的荷兰人范德瓦尔登的绝妙的结果。当时，他还是一个初出茅庐的青年，但现在已是一个知名的学者了。他的结果恰好是在哥庭根得到的，实际上，仅在我到达的前几天才得到，但是，我所遇到的几乎所有的数学家都津津有味地跟我谈论这个结果。

经过是这样的：当地的一个数学家（我忘记他的姓名）在自己的研究工作中碰到如下的问题：设全体自然数集以任意方式分成两部分（例如偶数与奇数，或素数与合数，或其它任意方式），那么，是否可以保证，至少在其中一部分中，有任意长的算术级数存在（在这儿及此后，所谓算术级数的长即指它的项数）？着手这一问题的人，开始时都感到相当简单，觉得它的肯定回答几乎是不证自明的。但真的要去证明时，却一筹莫展了。因为哥庭根的数学家和他们的外国朋友之间，有经常性的学术交流的传统，这个困难而迷人的问题，很快就成了大家津津乐道的话题。从年高望重的数学大师到低年级的

大学生，大家都在研究它。经过几星期的紧张奋战之后，终于被来哥庭根学习的荷兰青年范德瓦尔登成功地解决了。我认识他并亲耳听到他从容不迫的解答。这个解答是初等的，但远不是简单的；问题的内容很深刻，表面上的简单性仅是惑人的外衣。

不久前，M. A. 鲁科姆斯卡娅（明斯克人）为我提供一个新的、大为简单而又明显的证明。下面，承她允许，我将给你们叙述她的这个证明。

§ 2

从本质上说，范德瓦尔登证明的结果比原先要求的更多。首先，他假设自然数不是分成两类，而是分成任意 k 类（集合）；其次，为了保证至少有一类含给定（任意）长的算术级数，他指出，不一定要分全体自然数，而只要取某一段，这一段的长度 $n(k, l)$ 是 k 和 l 的函数，显然，在什么地方取这一段完全一样，只要它是 $n(k, l)$ 个连续的自然数。

因此，范德瓦尔登定理可表述如下：

设 k 和 l 是任意自然数，则存在自然数 $n(k, l)$，使得以任意方式分长为 $n(k, l)$ 的任意自然数段成 k 类（其中，可能有空类），则至少有一类，含有长为 l 的算术级数。

$l=2$ 时，这定理显然正确。因为，只要令 $n(k, 2)=k+1$，则分 $k+1$ 个数成 k 类，至少有一类含有不止一个数，而任两个数是长为 2 的算术级数，故定理得证。我们要就长 l，用数学归纳法证明定理。因此，以后我们将假设对某个数 $l\geqslant 2$ 和任意的 k，定理成立，并要证明对于数 $l+1$（同时，对于任意 k），定理也成立。

§ 3

根据我们的假设，对任意的自然数k，存在自然数$n(k, l)$，使得长为$n(k, l)$的任意自然数段用任意方式分成k类，至少有一类含有长为l的算术级数。我们的任务是证明：对任意自然数k，存在数$n(k, l+1)$。我们利用直接构造数$n(k, l+1)$的方法来解决这个问题。为此，令

$$q_0=1,\ n_0=n(k, l)。$$

然后用以下方式依次确定数$q_1, q_2, \cdots, n_1, n_2, \cdots$：如果对某个数$s>0$，已经确定了$q_{s-1}$和$n_{s-1}$，则我们令

$$q_s=2n_{s-1}q_{s-1},\ n_s=n(k^{q_s}, l) \quad (s=1, 2, \cdots)。 \tag{1}$$

显然，对任意$s\geqslant 0$，数n_s，q_s因此而确定。现在，我们将断言：可以取q_k为$n(k, l+1)$。为此，我们必须证明：如果长为q_k的自然数段以任意方式分成k类，则至少有一类含有长为$l+1$的算术级数。这一章的后半部分全是为了证实这一点。

为简便起见，以后令$l+1=l'$。

§ 4

设长为q_k的自然数段$\varDelta$以任意方式分成k类。两个数a和b属于同一类，我们称它们为同型，并记为$a\backsim b$。属于$\varDelta$的等长的两段$\delta(a, a+1, \cdots, a+r)$和$\delta'(a', a'+1, \cdots, a'+r)$，如果$a\backsim a'$，$a+1\backsim a'+1$，$\cdots$，$a+r\backsim a'+r$，则称这两段为同型，并记为$\delta\backsim\delta'$。显然，属于段$\varDelta$的数中，一切可能的不同的型数等于$k$，而对形为$(a, a+1)$的段（即长为2的段），一切可能的型数等于$k^2$，一般地，对长为$m$的段，一切可能的型

数等于 k^m（不言而喻，在 Δ 中，有些型可能实际上不存在）。

因 $q_k=2n_{k-1}q_{k-1}$（见(1)），则段 Δ 可视为 $2n_{k-1}$ 个长为 q_{k-1} 的段，象我们刚提到的，这样的段只能有 $k^{q_{k-1}}$ 个不同的型，但段 Δ 的左半部分有 n_{k-1} 个这样的段，由(1)，其中 $n_{k-1}=n(k^{q_{k-1}}, l)$，则依数 $n(k^{q_{k-1}}, l)$ 的意义，我们能断言：段 Δ 的左半部分含有 l 个长为 q_{k-1} 的同型段

$$\Delta_1,\ \Delta_2,\ \cdots,\ \Delta_l$$

组成的算术级数。在这儿，为简便起见，我们说等长的段 Δ_i 形成算术级数，指的是它们的第一个数形成算术级数。这个级数相邻两段的第一个数的差 d_1 称为级数 $\Delta_1,\ \Delta_2,\ \cdots,\ \Delta_l$ 的公差。不言而喻，这样两个相邻段的第二（或第三、第四等等）个数的差也等于 d_1。

现在，我们往这级数接上第 $l+1$ 项 $\Delta_{l'}$（记住，$l'=l+1$），它可能超出段 Δ 的左半部分，但无论如何，必整个地属于 Δ。则段 $\Delta_1,\ \Delta_2,\ \cdots,\ \Delta_l,\ \Delta_{l'}$ 表示长为 $l'=l+1$ 的算术级数且公差为 d_1，而每段长为 q_{k-1}，$\Delta_1,\ \Delta_2,\ \cdots,\ \Delta_l$ 是相互同型，至于最后一段 $\Delta_{l'}$ 的型则一无所知。至此完成了我们的第一步构造。在继续看下去之前，最好请你们先想一想，下一步该怎么办？

§ 5

现在，我们进行第二步。从刚才构造的段级数的前 l 项中任取一项，设取 Δ_{i_1}，其中 $1\leqslant i_1\leqslant l$，则 Δ_{i_1} 是长为 q_{k-1} 的段。与前面类似，因 $q_{k-1}=2n_{k-2}q_{k-2}$，故 Δ_{i_1} 的左半部分可分成 n_{k-2} 个长为 q_{k-2} 的段，但这样的段共有 $k^{q_{k-2}}$ 种不同的型，另一方面，由(1)，$n_{k-2}=n(k^{q_{k-2}}, l)$，因而，$\Delta_{i_1}$ 的左半部分含有长为 q_{k-2} 的由 l 个同型的段 $\Delta_{i_1i_2}(1\leqslant i_2\leqslant l)$ 组成的级数。设这个

级数的公差(即相邻两段第一个数的差)等于d_2。往这个级数接上第$l+1$项$\Delta_{i_1l'}$,它的型当然是一无所知,段$\Delta_{i_1l'}$可能超过段Δ_{i_1}的左半部分,但显然,整个地属于Δ_{i_1}。

现在把仅在其中一段Δ_{i_1}所做的构造全等地转移到所有其它段$\Delta_{i_1}(1\leqslant i_1\leqslant l')$。这样一来,得到带有两个下标的段集$\Delta_{i_1i_2}(1\leqslant i_1\leqslant l',\ 1\leqslant i_2\leqslant l')$。显然,下标不超过$l$的任意两段是同型的:

$$\Delta_{i_1i_2}\backsim\Delta_{i_1'i_2'}\quad(1\leqslant i_1,\ i_2,\ i_1',\ i_2'\leqslant l)。$$

现在,无疑你们已经明白,这个过程可以继续进行下去,直到k次。第一步构造的结果得到长为q_{k-1}的段,第二步得到长为q_{k-2}的段,等等,第k步得到长为$q_0=1$的段,即原先的段Δ的数;但我们仍然把这数表示成:

$$\Delta_{i_1i_2\cdots i_k}\quad(1\leqslant i_1,\ i_2,\ \cdots,\ i_k\leqslant l')。$$

这时,对$1\leqslant s\leqslant k$和$1\leqslant i_1,\ \cdots,\ i_s,\ i_1',\ \cdots,\ i_s'\leqslant l$,

$$\Delta_{i_1\cdots i_s}\backsim\Delta_{i_1'\cdots i_s'}。\tag{2}$$

现在,我们还要注意两点,它们对后面的论证是重要的。

1) 在关系式(2)中,如果$s<k$且$i_{s+1},\ i_{s+2},\ \cdots,\ i_k$是取自$1,\ 2,\ \cdots,\ l,\ l'$的任意下标,则数$\Delta_{i_1i_2\cdots i_si_{s+1}\cdots i_k}$在段$\Delta_{i_1\cdots i_s}$中所处的位置与数$\Delta_{i_1'i_2'\cdots i_s'i_{s+1}\cdots i_k}$在段$\Delta_{i_1'\cdots i_s'}$中所处的位置相同,由(2),这两段同型,故得

$$\Delta_{i_1i_2\cdots i_si_{s+1}\cdots i_k}\backsim\Delta_{i_1'i_2'\cdots i_s'i_{s+1}\cdots i_k},\tag{3}$$

其中

$$1\leqslant i_1,\ \cdots,\ i_s,\ i_1'\cdots,\ i_s'\leqslant l,$$
$$1\leqslant i_{s+1},\ i_{s+2},\ \cdots,\ i_k\leqslant l'\quad(1\leqslant s\leqslant k)。$$

2) 当$s\leqslant k$,$i_s'=i_s+1$时,显然,段$\Delta_{i_1\cdots i_{s-1}i_s}$和$\Delta_{i_1\cdots i_{s-1}i_s'}$是我们的第$s$步构造中相邻的两段。因此,无论下标$i_{s+1},\ \cdots,\ i_k$

如何，数 $\Delta_{i_1\cdots i_{s-1}i_si_{s+1}\cdots i_k}$ 和 $\Delta_{i_1\cdots i_{s-1}i'_si_{s+1}\cdots i_k}$ 在相邻两段中处于相同的位置，故（当 $i'_s=i_s+1$ 时）

$$\Delta_{i_1\cdots i_{s-1}i'_si_{s+1}\cdots i_k}-\Delta_{i_1\cdots i_{s-1}i_si_{s+1}\cdots i_k}=d_s。\tag{4}$$

§ 6

这时，已离我们的目的不远了。我们考虑段 Δ 的 $k+1$ 个数：

$$\begin{cases} a_0=\Delta_{l'l'l'\cdots l'}, \\ a_1=\Delta_{1l'l'\cdots l'}, \\ a_2=\Delta_{11l'\cdots l'}, \\ \cdots\cdots\cdots\cdots \\ a_k=\Delta_{111\cdots 1}。\end{cases}\tag{5}$$

因为段 Δ 中的数分为 k 类，而(5)中共有 $k+1$ 个数，故必有两个数属于同一类，设它们是 a_r 和 $a_s(r<s)$，即

$$\Delta_{\underbrace{1\cdots1}_{r}\underbrace{l'\cdots l'}_{k-r}}\backsim\Delta_{\underbrace{1\cdots1}_{s}\underbrace{l'\cdots l'}_{k-s}}\tag{6}$$

考虑 $l+1$ 个数

$$c_i=\Delta_{\underbrace{1\cdots1}_{r}\underbrace{i\cdots i}_{s-r}\underbrace{l'\cdots l'}_{k-s}}\quad(1\leqslant i\leqslant l'),\tag{7}$$

它们中的前 l 个数（即当 $i<l'$），由(3)知，属于同一类。至于最后一个数 $(i=l')$，由(6)知，它和第一个数同型。因此，(7)中所有的 $l+1$ 个数属于同一类。为证明我们的论断，只须验证这一些数形成算术级数，即差 $c_{i+1}-c_i(1\leqslant i\leqslant l)$ 与 i 无关。

为简便起见，令 $i+1=i'$，并令

$$c_{i,m}=\Delta_{\underbrace{1\cdots1}_{r}\underbrace{i'\cdots i'}_{m}\underbrace{i\cdots i}_{s-r-m}\underbrace{l'\cdots l'}_{k-s}}\quad(0\leqslant m\leqslant s-r),$$

则 $c_{i,0}=c_i$，$c_{i,s-r}=c_{i+1}$，故

$$c_{i+1}-c_i=\sum_{m=1}^{s-r}(c_{i,m}-c_{i,m-1})。$$

但由(4)，

$$c_{i,m}-c_{i,m-1}=\Delta_{\underbrace{1\cdots1}_{r}\underbrace{i'\cdots i'}_{m}\underbrace{i\cdots i}_{s-r-m}\underbrace{l'\cdots l'}_{k-s}}-\Delta_{\underbrace{1\cdots1}_{r}\underbrace{i'\cdots i'}_{m-1}\underbrace{i\cdots i}_{s-r-m+1}\underbrace{l'\cdots l'}_{k-s}}=d_{r+m},$$

故，差

$$c_{i+1}-c_i=d_{r+1}+d_{r+2}+\cdots+d_s$$

确实与 i 无关，这完全证明了我们的断言。

你们看，完全初等的证明有时是多么复杂！更有甚者，在下一章，你们将碰到同样初等但却复杂得多的问题。当然，我们完全不能认为，范德瓦尔登定理不能用更简单的方法证明，这方面的任何探索都是受欢迎的*。

* 格拉翰(Graham)和罗斯雪尔德(Rothschild)于1974年利用图论观点，提出了更为简单的证明。——译者注

第 二 章

朗道-斯尼列利曼猜测和曼恩定理

§ 1

也许你们已听说过著名的拉格朗日(Lagrange)定理：每个自然数是不多于四个平方数的和。也就是说：每个自然数或为另一个自然数的平方，或为两个、三个或四个自然数的平方和。稍后，我们将用某种不同的形式表述这个定理的内容。从 0 开始，写下平方数列：

$$0,\ 1,\ 4,\ 9,\ 16,\ 25,\ \cdots \qquad (Q)$$

这是一个整数列，用 Q 表示，与它完全相同的四个数列，分别记为 Q_1, Q_2, Q_3 和 Q_4。现在，从 Q_1 中任取数 a_1^2，从 Q_2 中任取数 a_2^2，从 Q_3 中任取数 a_3^2，而从 Q_4 中任取数 a_4^2，把这四个数加起来，得到和

$$n=a_1^2+a_2^2+a_3^2+a_4^2, \qquad (1)$$

它可能是

1) 零(如果 $a_1=a_2=a_3=a_4=0$)；

2) 自然数的平方(如果表达式 (1) 中的数 a_1, a_2, a_3, a_4 中有三个为 0，而第四个不为 0)；

3) 两个自然数的平方和(如果表达式(1) 中的数 a_1, a_2, a_3, a_4 中有两个为 0，而另两个不为 0)；

4) 三个自然数的平方和(如果表达式(1) 中的数 a_1, a_2, a_3, a_4 中有一个为 0，而其它三个不为 0)；

5) 四个自然数的平方和(如果表达式(1) 中的数 a_1, a_2, a_3, a_4 全不为 0)。

因此，得到的数 n 或为 0，或可表为不多于四个平方数的和的形式；显然，反过来，所有这样的自然数可用我们刚才描述的过程得到。

现在，从刚才给你们指出的过程得到的自然数(即分别取自数列 Q_1, Q_2, Q_3 和 Q_4 的四个数的和)，我们把它们依大小排成序列

$$0,\ n_1,\ n_2,\ n_3,\ \cdots \tag{A}$$

(这里，$0<n_1<n_2<n_3<\cdots$，为此，如果得到的数中，有些相等，则只取其一列入(A)中)。这时，拉格朗日定理只是断言数列(A)含有全体自然数，即 $n_1=1$, $n_2=2$, $n_3=3$ 等等。

现在，我们推广这个过程。设有 k 个从 0 开始的递增整数列：

$$0,\ a_1^{(1)},\ a_2^{(1)},\ \cdots,\ a_m^{(1)},\ \cdots, \qquad (A^{(1)})$$

$$0,\ a_1^{(2)},\ a_2^{(2)},\ \cdots,\ a_n^{(2)},\ \cdots, \qquad (A^{(2)})$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \qquad \cdots\cdots$$

$$0,\ a_1^{(k)},\ a_2^{(k)},\ \cdots,\ a_m^{(k)},\ \cdots。 \qquad (A^{(k)})$$

从每个数列 $A^{(i)}(1\leqslant i\leqslant k)$ 中任取一数，并把这 k 个数加起来，这样得到的数的全体排成新的无重复的递增数列，得

$$0,\ n_1,\ n_2,\ \cdots,\ n_m,\ \cdots, \tag{A}$$

则我们称它为已知数列 $A^{(1)}$, $A^{(2)}$, $\cdots$, $A^{(k)}$ 的和：

$$A=A^{(1)}+A^{(2)}+\cdots+A^{(k)}=\sum_{i=1}^{k}A^{(i)}。$$

那么，拉格朗日定理的内容是：和 $Q+Q+Q+Q$ 含全体自然数。

也许，你们知道著名的费尔马定理：和 $Q+Q$ 含有被 4

除余1的素数(即数5, 13, 17, 29, …)全体。也许你们也知道，著名的苏联学者维诺格拉多夫证明了以下的定理：用P表示由0和全体素数组成的数列

$$0, 2, 3, 5, 7, 11, 13, 17, \cdots \qquad (P)$$

则和$P+P+P$含有全体充分大的素数。许多最伟大的数学家曾为这定理进行了两百年无成效的奋斗*。

我在这儿介绍这些例子的唯一的十分简单的目的，是使你们熟悉数列和的概念，并表明：借助这个概念，数论的一些经典定理的阐述是多么方便和简单。

§ 2

毫无疑问，你们一定注意到，在上节的所有例子中，我们竭力要确立：一定个数的数列之和是完全或几乎完全含有另一数类(例如：全体自然数，充分大的素数等等)。在一切其它类似的问题中，查明已知数列的和是否在某种意义上稠密地分布在自然数列中，这恰是我们的研究目的。这时，经常谈到和含有全体自然数(如我们在第一个例子中看到的)。拉格朗日定理断言：四个数列Q的和含有全体自然数。一般地，如果k个同样的数列A的和含有全体自然数，则称数列A为自然数列的k阶基。这样一来，拉格朗日定理断言：平方数列Q是四阶基。稍后，我们将指明立方数列是九阶基。容易看出，一切k阶基同时也是$k+1$阶基。

在这些及许多其它的例子中，和的“密率”由被加数列的特殊性质——即这些数列具备的算术性质(它们或是平方，或

* 这定理实为著名的哥德巴赫猜想的减弱形式。——译者注

是素数，或是其它的性质)——所决定。著名的苏联学者Л.Г.斯尼列利曼在1930年首先提出这样的问题：数列和的密率在怎样的程度上只决定于被加数列的密率，而与它们的算术性质无关？这个问题不仅意义深远和饶有趣味，而且有助于处理一些经典问题。它给许多出色的研究提供了有力的工具，也有了丰富的文献资料。

为了能在这领域准确地提出问题并不加引号地书写词"密率"，我们必须首先约定应该用怎样的数来度量被研究数列的"密率"(恰如在物理学中，词"热"和"冷"得到准确的科学意义仅在能够量测温度之后)。

在我们研究各种问题中采用的"密率"，其十分方便的度量是斯尼列利曼提出的。设数列

$$0,\ a_1,\ a_2,\ \cdots,\ a_n,\ \cdots, \qquad (A)$$

如通常所要求的，所有的 a_n 是自然数，$a_n<a_{n+1}(n=1,2,\cdots)$，用 $A(n)$ 表示数列 (A) 中不超过 n 的自然数的个数(零不算在内)，则 $0\leqslant A(n)\leqslant n$，故

$$0\leqslant\frac{A(n)}{n}\leqslant 1。$$

显然，分数 $\frac{A(n)}{n}$ 对不同的 n 有不同的值，它可视为数列 (A) 在从1到 n 的自然数段间的一种平均密度。这些分数全体的最大下界，斯尼列利曼建议称之为数列 (A) 在全体自然数列中的"密率"。我们用 $d(A)$ 来表示它。

为了掌握这个概念的最简单的性质，我建议你们独立证明下列命题：

1) 如果 $a_1>1$(即数列 (A) 不含1)，则 $d(A)=0$。

2) 如果 $a_n=1+r(n-1)$(即数列 (A) 从 a_1 开始是首项

为1，公差为 r 的算术级数)，则

$$d(A)=\frac{1}{r}。$$

3) 一切几何级数的密率是0。

4) 平方数列的密率是0。

5) 为了数列(A)含有全体自然数$(a_n=n,\ n=1,\ 2,\ \cdots)$，必须且只须 $d(A)=1$。

6) 如果 $d(A)=0$ 且(A)含数1，则对任意的 $\varepsilon>0$，可以找到充分大的数 N，使得

$$A(N)<\varepsilon N。$$

如果你们证明了所有这一些命题，那么，就能熟悉密率的概念并能应用它了。现在，我希望你们还能证明下面这个虽则简单，但十分著名的斯尼列利曼引理：

$$d(A+B)\geqslant d(A)+d(B)-d(A)d(B)。\qquad (2)$$

这个不等式的意思可理解为：**任意两个数列的和的密率不小于它们密率的和减去密率的积**。这个"斯尼列利曼不等式"，对于用被加数列的密率来估计和的密率是第一个意义深远的工具。设 $A(n)$ 表示在数列 A 中不超过 n 的自然数的个数，$B(n)$是数列 B 中不超过 n 的自然数的个数，为简便起见，令 $d(A)=\alpha,\ d(B)=\beta,\ A+B=C,\ d(C)=\gamma$。自然数段$(1,\ n)$中含 $A(n)$个数在数列(A)中，它们也都在数列 C 中。设 a_k 和 a_{k+1} 是这些数中相邻的两个数，在它们之间有 $a_{k+1}-a_k-1=l$ 个数不在 A 中，即数

$$a_k+1,\ a_k+2,\ \cdots,\ a_k+l=a_{k+1}-1。$$

但它们有一些在 C 中，例如形为 a_k+r 的一切数，其中 r 在 B 中（我们将简便地写为 $r\in B$）。但是最后这种形式的数恰与段$(1,\ l)$中含数列 B 的数有相同的个数即 $B(l)$，因此，含在数

列 A 中相邻两数间长为 l 的一切段，含 C 的数的个数不少于 $B(l)$，故得段 $(1, n)$ 中在 C 中的数的数目 $C(n)$ 不少于

$$A(n)+\sum B(l),$$

和号取遍上述一切段。依密率定义，$B(l)\geqslant\beta l$，故

$$C(n)\geqslant A(n)+\beta\sum l=A(n)+\beta\{n-A(n)\},$$

因为 $\sum l$ 是端点只在 A 中的各段长的和，即段 $(1, n)$ 中不在 A 的数的个数 $n-A(n)$。但 $A(n)\geqslant\alpha n$，故

$$C(n)\geqslant A(n)(1-\beta)+\beta n\geqslant\alpha n(1-\beta)+\beta n。$$

由此得
$$\frac{C(n)}{n}\geqslant\alpha+\beta-\alpha\beta。$$

因为这个不等式对任意自然数 n 成立，故

$$\gamma=d(C)\geqslant\alpha+\beta-\alpha\beta,$$

这就是所要证明的。

斯尼列利曼不等式(2)可写成等价形式

$$1-d(A+B)\leqslant\{1-d(A)\}\{1-d(B)\},$$

由这形式不难推广到任意个被加项的情况：

$$1-d(A_1+A_2+\cdots+A_k)\leqslant\prod_{i=1}^{k}\{1-d(A_i)\}。$$

其证明可用简单的归纳法，你们不难证实它。如果把最后一个不等式写成

$$d(A_1+A_2+\cdots+A_k)\geqslant1-\prod_{i=1}^{k}\{1-d(A_i)\}, \qquad (3)$$

则可用被加项密率来估计和的密率。斯尼列利曼从他的初等不等式出发，推出一系列十分著名的结论。首先是以下重要的定理：

一切正密率的集合是自然数列的基。

易言之，如果 $\alpha=d(A)>0$，则足够多的数列 A 的和含全部自然数。这个定理的证明是如此简单，尽管它稍微偏离我

们原先的问题，我还是想给你们讲讲它的证明。

为简便计，我们用 A_k 表示 k 个与 A 相同的数列的和，则由不等式(3)，

$$d(A_k)\geqslant 1-(1-\alpha)^k;$$

因 $\alpha>0$，当 k 足够大时，

$$d(A_k)>\frac{1}{2}。\qquad (4)$$

现在，不难证明数列 A_{2k} 含全部自然数。这可从以下一般的命题推出。

引理 如果 $A(n)+B(n)>n-1$，则 n 在 $A+B$ 中。

实际上，如果 n 在 A 或在 B 中，则得证。故我们可以假设 n 不在 A 也不在 B 中，则

$$A(n)=A(n-1),\ B(n)=B(n-1),$$

因而

$$A(n-1)+B(n-1)>n-1。$$

设 $a_1, a_2, \cdots, a_r$ 和 $b_1, b_2, \cdots, b_s$ 分别是在 A 中和 B 中属于段 $(1, n-1)$ 的数，则所有的数

$$a_1, a_2, \cdots, a_r,$$

$$n-b_1, n-b_2, \cdots, n-b_s$$

在段 $(1, n-1)$ 中，它们的数目是 $r+s=A(n-1)+B(n-1)>n-1$，故上行的数至少有一个等于下行的某一个数，设 $a_i=n-b_k$，则 $n=a_i+b_k$，即 n 在 $A+B$ 中。

现在回到我们原先的讨论中来。因(4)，对任意的 n，得

$$A_k(n)>\frac{1}{2}n>\frac{n-1}{2},$$

这表明

$$A_k(n)+A_k(n)>n-1,$$

故由刚才证明的引理，n 在 $A_k+A_k=A_{2k}$ 中；但 n 是任意的自然数，故我们的定理证毕。

在斯尼列利曼的论文中，由这简单的定理得到一系列重要的推论。他第一个证明了：由1和全体素数组成的数列是自然数的基。诚然，这个数列P，正如欧拉证明的，密率是0，因而不能直接应用刚证明的定理，但斯尼列利曼成功地证明了$P+P$有正密率，即$P+P$是基，故P也是基。由此立即得出：对足够大的k，除1以外的自然数可表为不多于k个素数的和的形式。当时(1930年)，这是个重大的成果，从而引起了科学界的极大兴趣。正如我在这节开头跟你们说到的，因为维诺格拉多夫的出色研究，在这方面已有更深入的结果了。

§ 3

前面的那些，其目的是尽可能快地把你们引导到数论中独特的和有趣的领域中。斯尼列利曼的著作开创了这领域的研究。但是，这个领域的一个特殊问题是本章的直接目的。我现在转而阐述这个问题。

1931年秋，斯尼列利曼从国外出差回来，报告他在哥庭根和朗道的谈话，顺便说到他们发现了如下有趣的事实：在他们能够想到的一切具体例子中，我们在§2中导出的不等式

$$d(A+B)\geqslant d(A)+d(B)-d(A)d(B),$$

可以用更强(也更简单)的不等式

$$d(A+B)\geqslant d(A)+d(B) \tag{5}$$

代替，即和的密率永远不小于被加项的密率之和(在这儿当然要求$d(A)+d(B)\leqslant 1$)。自然的，他们因而猜想不等式(5)是一般规律，但一着手证明这个猜想，开始并未成功。这立即成为很显然的事，如果他们的猜想是正确的，其证明方法一定是

很复杂的。我们也将注意到，如果猜想的不等式(5)，实际上是一般的规律，则借助于数学归纳法可立即推广到任意个被加项的情况，即当条件 $\sum_{i=1}^{k} d(A_i) \leqslant 1$ 成立时，则不等式

$$d(\sum_{i=1}^{k} A_i) \geqslant \sum_{i=1}^{k} d(A_i) \tag{6}$$

也成立。

这个问题，因为它的简单和精致，同时，也因为它的初等性及解决它的困难性，自然引起了研究者的注意。当时，我自己也被它迷恋上了，并为它放弃了其它所有的研究。经过几个月的紧张努力之后，在1932年初，我证明不等式(5)在重要的特殊情况 $d(A)=d(B)$ 下成立（应该承认，这种情况是最重要的，因为在许多具体问题中，所有被加项都是一样的）。同时，我证明了一般的不等式(6)当 $d(A_1)=d(A_2)=\cdots=d(A_k)$ 时成立（不难看出，这结果不能从上一个结果用简单的归纳法得到，而要求单独证明）。我用的方法完全是初等的，但很繁。后来，我把证明略为简化一些。

不管怎样，这些都仅是特殊情况。我一直以为，我的方法经某些适当而巧妙的改进，就能完全解决这个问题，但是，我在这方面的一切努力没有任何收效。

当时，我的著作一发表，就吸引了世界上相当多的研究者对朗道-斯尼列利曼猜想的注意，得到许多并非十分有意义的特殊结果，产生了一系列文献。有一些作者把问题从自然数领域扩充到其它领域。总之，问题变得"时髦"了，科学界为它提供了奖金。在1935年，我的英国朋友写信告诉我说：英国至少有一半的数学家把他们的日常事务搁置一旁，试图解决这个问题。朗道在论述堆垒数论的最新成就的书中写道，希望"读者把这个问题记在心里"。但它显得很顽固，最能干的

研究者经过了好几年的努力也攻克不下它。直到1942年末，年轻的美国数学家曼恩才攻克，他找到(5)(从而也得到(6))的完全的证明。他的方法完全是初等的，依风格而论，接近于我的方法，但是基于完全不同的另一种思想。其证明很繁很长，在此，我不想给你们介绍这个证明。在1943年，阿亭和谢尔克发表了一个新的证明，它完全是基于另一种思想，尽管同样也是初等的，但较为易懂且简短得多。这就是我要向你们介绍的证明，也是我写这一章的目的。它是以下各节的内容。

§ 4

设 A 和 B 是两个数列，令 $A+B=C$，$A(n)$ 和 $d(A)$ 等有通常的意义。记住，我们的数列都是从0开始，而计算 $A(n)$，$B(n)$，$C(n)$，只考虑到这些数列中的自然数。我们要证明：只要 $d(A)+d(B)\leqslant 1$，则不等式

$$d(C)\geqslant d(A)+d(B) \tag{7}$$

成立。以后，为简便计，设 $d(A)=\alpha$，$d(B)=\beta$。

基本引理 *对任意自然数 n，存在整数 m $(1\leqslant m\leqslant n)$ 使得*

$$C(n)-C(n-m)\geqslant(\alpha+\beta)m。$$

也就是说，在段 $(1, n)$ 中存在"末端" $(n-m+1, n)$，使数列 C 在这末端上的平均密度不小于 $\alpha+\beta$。

现在，我们面临着两个问题：1). 证明基本引理；2). 从基本引理推出不等式(7)。其中第二个问题比第一个问题简单，因而，首先解决它。

设基本引理成立，则在段 $(1, n)$ 的某末端 $(n-m+1, n)$，

数列C的平均密度不小于$\alpha+\beta$。但在段$(1, n-m)$上，由于基本引理，又有某末端$(n-m-m'+1, n-m)$，数列C在它上面的平均密度不小于$\alpha+\beta$。显然，如此继续下去，经过有限次，段$(1, n)$分成有限小段，其每一段上，C的平均密度不小于$\alpha+\beta$，故在整段$(1, n)$上，数列C的平均密度不小于$\alpha+\beta$，因n是任意的，故有

$$d(C)\geqslant\alpha+\beta,$$

这正是所要求证的。

因而，问题归结为证明基本引理。为此，我们要用较长的篇幅和较复杂的技巧。

§5 正规数列

下面，我们将认为数n是固定的，而所研究的数列都是由0和段$(1, n)$中的某些数组成。数列H认为是正规的，如果它具有下列性质：对段$(1, n)$中不属H的任意数f和f'，数$f+f'-n$也不属H（不排除$f=f'$的情况）。

如果数n属于数列C，则

$$C(n)-C(n-1)=1\geqslant(\alpha+\beta)\times 1,$$

故基本引理为真（取$m=1$）。因此，以后，请记住，我们将假设n不在C中。

首先，我们不难验证，当C是正规数列时，基本引理成立。实际上，用m表示不在C中的最小自然数（因为依假设，n不在C中，故$m\leqslant n$）。设s是在$n-m$和n之间的任意数，$n-m<s<n$，则$0<s+m-n<m$，故$s\in C$。实际上，若不然，则由C的正规性，数$s+m-n$不在C中，但我们刚才指出，这个数小于m，而依假设，m是不在C中的最小自然数。

这样一来，段 $n-m<s<n$ 中的所有数 s 在 C 中，故

$$C(n)-C(n-m)=m-1。$$

另一方面，因 m 不在 $C=A+B$ 中，由§2的引理，$A(m)+B(m)\leqslant m-1$，故

$$C(n)-C(n-m)\geqslant A(m)+B(m)\geqslant(\alpha+\beta)m, \tag{8}$$

即得基本引理。

§6 典式扩张

现在讨论 $C=A+B$ 不具有正规性的情况。这时，我们将依一定规则，由不在 B 中的某些数组成一个新的集合，并把它附到 B 上，得到扩张集 B_1，显然，$A+B_1=C_1$ 是 C 的扩张集。如上指出，B 和 C 的扩张集(集合 A 不变)是依唯一确定的方式定义，它们当且仅当 C 不是正规时，才可能产生。我们将称这扩张为 B 和 C 的典式扩张。最后，我们将导出典式扩张的一些重要性质，并借助它完成基本引理的证明。

先给集合 B 和 C 的典式扩张以精确的定义。如果 C 不是正规的，则段 $(0, n)$ 中存在数 c 和 c'，使得

$$c\overline{\in}C,\ c'\overline{\in}C,\ c+c'-n\in C。$$

因 $C=A+B$，故有

$$c+c'-n=a+b\quad(a\in A,\ b\in B)。\tag{9}$$

设 β_0 是集合 B 中能在等式(9)中起数 b 作用的最小数。用另一句话说，β_0 是最小的数 $b\in B$，使得在段 $(0, n)$ 中的数 $c\overline{\in}C$，$c'\overline{\in}C$，$a\in A$，经过适当的选择，等式(9)成立。我们称这数为扩张的基。

这样，方程

$$c+c'-n=a+\beta_0 \tag{10}$$

一定有解 c, c', a，它们满足条件

$$c\overline{\in}C,\ c'\overline{\in}C,\ a\in A,$$

同时，这三个数都属于段 $(0,\ n)$。满足方程(10)及上述条件的数 c 和 c'，形成集合 C^*，显然，C 和 C^* 没有公共元素，其并(即或在 C 中，或在 C^* 中的数全体)

$$C\cup C^*=C_1$$

称为 C 的典式扩张。

现在研究表达式 β_0+n-c，如果 c 跑遍刚刚构造的集合 C^* 的所有数，则这表达式的值的全体构成某个集合 B^*。因(10)，每个这样的数 $\beta_0+n-c(c\in C^*)$ 是形如 $c'-a$，其中 $c'\in C^*,\ a\in A$。

设 b^* 是 B^* 中任意的数，因有形式 β_0+n-c，故它 $\geqslant\beta_0\geqslant 0$，又因有形式 $c'-a(c'\in C^*,\ a\in A)$ 故它 $\leqslant c'\leqslant n$，因而，集合 B^* 的数在 $(0,\ n)$ 中。此外，如果 $b^*\in B^*$，则 $b^*\overline{\in}B$。因为否则，由 $b^*=c'-a$ 得 $c'=a+b^*\in A+B=C$，矛盾。这样一来，B^* 在段 $(0,\ n)$ 中与 B 没有公共元素，令

$$B\cup B^*=B_1,$$

我们称 B_1 为集合 B 的典式扩张。

我们首先要验证：

$$A+B_1=C_1。$$

先设 $a\in A,\ b_1\in B_1$，我们要证明 $a+b_1\in C_1$。由 $b_1\in B_1$ 得到：或 $b_1\in B$，或 $b_1\in B^*$。如果 $b_1\in B$，则 $a+b_1\in A+B=C\subset C_1$；如果 $b_1\in B^*$，则 $a+b_1$ 或在 C 中故在 C_1 中，或 $a+b_1\overline{\in}C$。这时(因 b_1 是 B^* 的元素，有形式 $\beta_0+n-c',\ c'\overline{\in}C$)得

$$c=a+b_1=a+\beta_0+n-c'\overline{\in}C。$$

故 $$c+c'-n=a+\beta_0\in A+B=C,$$

而 $c\overline{\in}C,\ c'\overline{\in}C$。依集合 C^* 的定义得

$$c=a+b_1\in C^*\subset C_1,$$

这样，我们证明了 $A+B_1\subset C_1$。

为了证明相反的关系，设 $c\in C_1$，由此得到：或 $c\in C$，或 $c\in C^*$。如果 $c\in C$，则 $c=a+b$，$a\in A$，$b\in B\subset B_1$；如果 $c\in C^*$，则我们已知，对某个 $a\in A$，数 $b^*=c-a$ 在 B^* 中，故 $c=a+b^*\in A+B^*\subset A+B_1$。因此 $C_1\subset A+B_1$。又因前面已证 $A+B_1\subset C_1$，说明 $C_1=A+B_1$。

现在，我要提醒你们，照我们原先的假设，$n\overline{\in} C$。不难看出（这对后面的论证是重要的），数 n 也不在扩张集 C_1 中。实际上，如果 $n\in C^*$，依 C^* 的定义，在关系式(10)中，令 $c'=n$，得 $c=a+\beta_0\in A+B=C$，但依关系式(10)的意义，$c\overline{\in} C$。

如果扩张集 C_1 还不是正规的，则由 $A+B_1=C_1$ 和 $n\overline{\in} C_1$，集合类 A，B_1，C_1 有 A，B，C 一样的性质，可进行新的典式扩张。找到这个扩张的新基，类似于上面的定义，补充集合 B_1^*，C_1^*，并令

$$B_1\cup B_1^*=B_2,\quad C_1\cup C_1^*=C_2。$$

同理可证 $A+B_2=C_2$ 和 $n\overline{\in} C_2$。这过程显然可以继续下去，直到扩张集 C_h 是正规的。这种情况必定出现的原因在于：每一次扩张时，我们都在集合 B_μ 和 C_μ 中加进段 $(0,\ n)$ 中不属于 B_μ 和 C_μ 的新数。

因此，得到有限集合列

$$B=B_0\subset B_1\subset\cdots\subset B_h,$$

$$C=C_0\subset C_1\subset\cdots\subset C_h,$$

同时，一切 $B_{\mu+1}$（相应的 $C_{\mu+1}$）含有不在 $B_\mu(C_\mu)$ 中的数，这些数组成集合 $B_\mu^*(C_\mu^*)$，使

$$B_{\mu+1}=B_\mu\cup B_\mu^*,\quad C_{\mu+1}=C_\mu\cup C_\mu^*(0\leqslant\mu\leqslant h-1)。$$

我们用 β_μ 表示从 $(B_\mu,\ C_\mu)$ 到 $(B_{\mu+1},\ C_{\mu+1})$ 扩张的基。而且

$$A+B_\mu=C_\mu,\ n\overline{\in}C_\mu\quad(0\leqslant\mu\leqslant h)。$$

最后，C_h 是正规的，而集合 $C_\mu(0\leqslant\mu\leqslant h-1)$ 则不是正规的。

§7 典式扩张的性质

以后所需的关于典式扩张的性质，我们用三个引理表述出来并给以证明。证明基本引理只用到最后一个引理，而引理一和引理二仅在证明引理三时用到。

引理一 $\beta_\mu>\beta_{\mu-1}(1\leqslant\mu\leqslant h-1)$，即典式扩张列的基是一个递增数列。

实际上，因 $\beta_\mu\in B_\mu=B_{\mu-1}\cup B_{\mu-1}^*$，则或 $\beta_\mu\in B_{\mu-1}^*$，这时，β_μ 有形式

$$\beta_\mu=\beta_{\mu-1}+n-c,$$

其中，$c\in C_{\mu-1}^*\subset C_\mu$，因 $c<n$，故 $\beta_\mu>\beta_{\mu-1}$，引理一得证；或 $\beta_\mu\in B_{\mu-1}$，这时，依 β_μ 的定义，存在 $a\in A$，$c\overline{\in}C_\mu$，$c'\overline{\in}C_\mu$，具有关系：

$$c+c'-n=a+\beta_\mu\in C_\mu,$$

但因 $\beta_\mu\in B_{\mu-1}$，故

$$c+c'-n=a+\beta_\mu\in A+B_{\mu-1}=C_{\mu-1},\tag{11}$$

而 $c\overline{\in}C_{\mu-1}$，$c'\overline{\in}C_{\mu-1}$，由 $\beta_{\mu-1}$ 的最小性得 $\beta_\mu\geqslant\beta_{\mu-1}$，但如 $\beta_\mu=\beta_{\mu-1}$，依集合 $C_{\mu-1}^*$ 的定义和式(11)得

$$c\in C_{\mu-1}^*\subset C_\mu,\ c'\in C_{\mu-1}^*\subset C_\mu,$$

都不真，故 $\beta_\mu>\beta_{\mu-1}$。

以后，我们将用 m 表示不在 C_h 中的最小正整数。

引理二 如果 $c\in C_\mu^*$，$0\leqslant\mu\leqslant h-1$，且 $n-m<c<n$，则 $c>n-m+\beta_\mu$，即在段 $n-m<c<n$ 中集合 C_μ^* 的所有数含在这段的一部分中，这部分用不等式 $n-m+\beta_\mu<c<n$ 表示。

只须要证明不等式

$$c+m-n>\beta_\mu。$$

从 $n-m<c<n$，得

$$0<m+c-n<m,$$

由此，依 m 的定义，得

$$m+c-n\in C_h。$$

但

$$C_h=C_\mu\cup C_\mu^*\cup C_{\mu+1}^*\cup\cdots\cup C_{h-1}^*。$$

因此，下面分两种情况讨论。

1) 如果 $m+c-n\in C_\mu$，则

$$m+c-n=a+b_\mu,\ a\in A,\ b_\mu\in B_\mu。$$

但 $m\overline{\in}C_\mu$，$c\overline{\in}C_\mu$(后者因 $c\in C_\mu^*$)。又因 β_μ 的最小性，应有 $b_\mu\geqslant\beta_\mu$。但当 $b_\mu=\beta_\mu$，依集合 C_μ^* 的定义，$m\in C_\mu^*$，这因 $C_\mu^*\subset C_{\mu+1}\subset C_h$ 和 $m\overline{\in}C_h$ 而不真。因此，$b_\mu>\beta_\mu$，故

$$m+c-n=a+b_\mu\geqslant b_\mu>\beta_\mu,$$

引理二得证。

2) 如果 $c'=m+c-n\in C_\nu^*(\mu\leqslant\nu\leqslant h-1)$，则依集合 C_ν^* 定义，c' 满足方程(10)，

$$c'-a=\beta_\nu+n-c'',$$

其中 $a\in A$，$c''\in C_\nu^*$。由此 $c'\geqslant c'-a>\beta_\nu\geqslant\beta_\mu$(后者由引理一得出)，仍证得引理二。

引理三 $C_\mu^*(n)-C_\mu^*(n-m)=B_\mu^*(m-1)\ (0\leqslant\mu\leqslant h-1)$。即在段 $n-m<c<n$ 中，$c\in C_\mu^*$ 的个数恰等于在(等长)段 $0<b<m$ 中 $b\in B_\mu^*$ 的个数。

研究关系式

$$b=\beta_\mu+n-c。\tag{12}$$

依集合 B_μ^* 和 C_μ^* 的定义，由 $c\in C_\mu^*$ 得 $b\in B_\mu^*$，反之亦然。且如 $n-m+\beta_\mu<c<n$，则 $\beta_\mu<b<m$，反之亦然。故

$$C_\mu^*(n) - C_\mu^*(n - m + \beta_\mu) = B_\mu^*(m-1) - B_\mu^*(\beta_\mu)。$$

但由引理二，$C_\mu^*(n-m+\beta_\mu) = C_\mu^*(n-m)$。另一方面，由(12)表示的一切 $b \in B_\mu^*$，因 $c < n$，故大于 β_μ，因此 $B_\mu^*(\beta_\mu) = 0$，则得

$$C_\mu^*(n) - C_\mu^*(n-m) = B_\mu^*(m-1),$$

这正是所要证明的。

§8 基本引理的证明

由§5的结果和刚刚证明的引理三，现在，我们能够很容易地证明基本引理。

对数列 A, B_h, C_h 应用形为不等式(8)的结果（因 C_h 的正则性，这是容许的），得

$$C_h(n) - C_h(n-m) \geqslant A(m) + B_h(m), \tag{13}$$

其中 m 是不在 C_h 中的最小正整数。显然 $m \overline{\in} A$ 且 $m \overline{\in} B_h$。故 $A(m)$ 和 $B_h(m)$ 分别可写成 $A(m-1)$ 和 $B_h(m-1)$。

因为在每组并

$$C_h = C \cup C^* \cup C_1^* \cup \cdots \cup C_{h-1}^*,$$

$$B_h = B \cup B^* \cup B_1^* \cup \cdots \cup B_{h-1}^*,$$

中的集合两两之间没有公共元素，故

$$C_h(n) - C_h(n-m) = C(n) - C(n-m) + \sum_{\mu=0}^{h-1} \{C_\mu^*(n) - C_\mu^*(n-m)\},$$

$$B_h(m) = B_h(m-1) = B(m-1) + \sum_{\mu=0}^{h-1} B_\mu^*(m-1),$$

在此，当然 $C_0^* = C^*$, $B_0^* = B^*$。由(13)得

$$C(n) - C(n-m) + \sum_{\mu=0}^{h-1} \{C_\mu^*(n) - C_\mu^*(n-m)\}$$

$$\geqslant A(m)+B(m-1)+\sum_{\mu=0}^{h-1} B_{\mu}^{*}(m-1)。$$

但是，由引理三

$$C_{\mu}^{*}(n)-C_{\mu}^{*}(n-m)=B_{\mu}^{*}(m-1) \quad (0\leqslant\mu\leqslant h-1),$$

故得

$$C(n)-C(n-m)\geqslant A(m)+B(m-1)$$
$$=A(m)+B(m)\geqslant(\alpha+\beta)m,$$

由此证明了基本引理。

从而，正如我们在§4看到的，曼恩定理得到完全的证明。这个定理是标志着堆垒数论诞生的具有决定性意义的基本定理。

阿亭和谢尔克的构造，难道不是一个辉煌灿烂的杰作？其构造的奥妙完美和极端初等高度地揉合，特别使我迷恋。

第 三 章

华林问题的初等证明

§1

回想上一章开头我跟你们谈到的拉格朗日定理：每个自然数可表为不多于四个平方数的和。我还跟你们谈到，这个定理完全可用另一种术语表述：数列

$$0,\ 1^2,\ 2^2,\ \cdots,\ k^2,\ \cdots \qquad (A_2)$$

自相加四次，得到的数列含全体自然数。或更简单地说：数列(A_2)是（自然数列的）四阶基。我还提到，稍后，我们会知道，立方数列

$$0,\ 1^3,\ 2^3,\ \cdots,\ k^3,\ \cdots \qquad (A_3)$$

是九阶基。所有这些事实，自然的，引出命题：对任何自然数n，数列

$$0,\ 1^n,\ 2^n,\ \cdots,\ k^n,\ \cdots \qquad (A_n)$$

是某阶基（当然，阶数与n有关）。实际上，这个猜测早在十八世纪已被华林提出，但解决它很不容易，到本世纪初（1907年）希尔伯特才完全证明了华林猜测的正确性。希尔伯特的证明不但在形式关系上很累赘，并需要很繁杂的解析理论（多重积分），而且，在思想方法上也很难理解。法国的著名数学家庞加莱在评论希尔伯特的科学创造时写道：如果产生这个证明的基本动机在某个时候被理解的话，那么，大量的数论成果也许将如雪花般飘来。在某种意义上，他的话是对的。10～15

年以后，英国的哈代和李特伍德与苏联的维诺格拉多夫给希尔伯特定理以新的证明。和希尔伯特的证明一样，这个证明是解析的，形式上很繁，但它的逻辑方法明显，思路简单，因而显得更为优越，它不能期望有任何的改进了。实际上，两个证明方法都成了新的数论定理的强大源泉。

但是，当我们的科学是涉及到如华林问题这样完全初等的问题时，那么，就应该给它一个解答，这个解答不必利用超出初等数论范围的概念和方法。阐述华林问题的这种初等证明，是我要和你们讨论的第三个问题。希尔伯特定理的完全初等的证明，到 1942 年才被年轻的苏联学者 Ю.В. 林尼克找到。

现在，你们已经明白，初等并不意味着简单。林尼克找到的华林问题的初等证法，正如你们将要看到的，很不简单。为了理解并掌握它，你们必须做出不少努力。我尽可能叙述得使你们易于理解。但你们要记住，在数学中(也许，在所有其它学科中)，掌握一切有价值、有意义的东西都须要紧张地劳动。

在林尼克的证明中，我在第二章给你们介绍的斯尼列利曼的思想起了十分本质的作用。回想一下(那时，我简单地说到)，斯尼列利曼为证明著名的定理：由 0，1 和所有素数组成的数列 P 是自然数的基，他指出，数列 $P+P$ 有正密率，因为由我们在第二章 §2 证明的一般的斯尼列利曼定理，一切正密率的数列是自然数的基，由此立即可得所要求的结论。这同样也是林尼克用以证明希尔伯特定理的基础。一切都归结为证明足够多的数列 (A_n) 之和是正密率数列，正如刚才所述，由此可认为，希尔伯特定理由一般的斯尼列利曼定理而得证。

§2 基本引理

依第二章的规则，把 k 个数列(A_n)相加，显然，得到的数列 $A_n^{(k)}$ 含 0 和一切这样的数，它们可表为至多 k 个形为 x^n 的数的和，其中 x 是任意的自然数。换言之，数 $\hat{m}$ 属于数列 $A_n^{(k)}$，仅当方程

$$x_1^n+x_2^n+\cdots+x_k^n=\hat{m}^* \tag{1}$$

在非负整数 $x_i(1\leqslant i\leqslant k)$ 的范围内有解。正如在§1所说明的，我们的目的是证明对足够大的 k，数列 $A_n^{(k)}$ 有正密率。

一般地说，对给定的 k 和 $\hat{m}$，方程(1)可能不止一个解。以后，我们用 $r_k(\hat{m})$表示其解数，即满足方程(1)的非负整数 $x_1, x_2, \cdots, x_k$ 的组数。显然，数 $\hat{m}$ 在 $A_n^{(k)}$ 中，当且仅当 $r_k(\hat{m})>0$。

下面，我们将把数 n 视为固定的，因此，仅与 n 有关的数认为是常数。这样的常数用字母 C 或 $C(n)$表示，而且，这样的常数 C 在讨论过程的不同地方可能有不同的值，只要求它们都是常数。常数的这种书写方法，实际上将简化证明的记号。

基本引理 存在只与 n 有关的自然数 $k=k(n)$ 和常数 C，使得对任意自然数 N

$$r_k(\hat{m})<CN^{\frac{k}{n}-1} \quad (1\leqslant\hat{m}\leqslant N)。 \tag{2}$$

象上一章，由此产生了两个问题。首先，证明基本引理；

---

* 原文中，符号 m 的用法较为混乱，有时在同一式中代表两种不同意思(如§7中开头的几个估计式)，此处用 $\hat{m}$ 代替原文的 m，以便区别。后面尚有多处，不另加注。——译者注

其次，由基本引理推出所要求的论断：数列 $A_n^{(k)}$ 有正密率。同样的，第二个问题比第一个问题大为容易，我们先证明它。

由数 $r_k(\hat{m})$ 的定义，立即可得和

$$r_k(0)+r_k(1)+\cdots+r_k(N)=R_k(N)$$

是由 k 个非负整数 $(x_1, x_2, \cdots, x_k)$ 组成，满足

$$x_1^n+x_2^n+\cdots+x_k^n\leqslant N \tag{3}$$

的数组数目。显然，当

$$0\leqslant x_i\leqslant\left(\frac{N}{k}\right)^{\frac{1}{n}}\quad(1\leqslant i\leqslant k)$$

时，满足上述要求，又对满足这个不等式的每个 x_i，有多于 $\left(\frac{N}{k}\right)^{\frac{1}{n}}$ 种不同的选择 $\left(x_i=0,\ 1,\ \cdots,\ \left[\left(\frac{N}{k}\right)^{\frac{1}{n}}\right]\right)$，$k$ 个（即数 x_1, $x_2,\cdots, x_k$ 的）选择能以任意方式组合，故这样的数组 $x_i(1\leqslant i\leqslant k)$ 有多于 $\left(\frac{N}{k}\right)^{\frac{k}{n}}$ 种不同选择，每种选择都满足条件(3)，这说明

$$R_k(N)\geqslant\left(\frac{N}{k}\right)^{\frac{k}{n}}。\tag{4}$$

依假设，基本引理正确，故对任意 N，不等式(2)成立。现在要证明不等式(2)和(4)只有在数列 $A_n^{(k)}$ 有正密率时，才有可能同时成立。下面讨论的思想很简单：在和 $R_k(N)$ 中，当 $\hat{m}$ 属于 $A_n^{(k)}$，被加项 $r_k(\hat{m})$ 才不等于 0，如 $A_n^{(k)}$ 的密率为 0，则对大的 N，被加项的项数相对而言比较小，又因(2)，每一项不可能很大，故其和 $R_k(N)$ 相对而言比较小，但由(4)，它又应足够大。

为此只须做一些演算。如果 $d(A_n^{(k)})=0$，则对任意小的 $\varepsilon>0$，有适当的 N，使不等式

$$A_n^{(k)}(N)<\varepsilon N$$

成立。在此可选 N 为充分大的数，这是由于 $A_n^{(k)}$（对任何 k）含数 1（回忆下你们自己证明的第二章 §2 的习题 6）。

利用估计式(2)得

$$R_k(N)=\sum_{m=0}^{N} r_k(m)=r_k(0)+\sum_{m=1}^{N} r_k(m)$$

$$<1+CN^{\frac{k}{N}-1}A_n^{(k)}(N)<1+C\varepsilon N^{\frac{k}{N}},$$

故对充分大的 N,

$$R_k(N)<2C\varepsilon N^{\frac{k}{n}}。$$

因对充分小的 ε,

$$2C\varepsilon<\left(\frac{1}{k}\right)^{\frac{k}{n}},$$

故得

$$R_k(N)<\left(\frac{N}{k}\right)^{\frac{k}{n}}$$

与(4)矛盾。故必有

$$d(A_n^{(k)})>0。$$

这正如上面所指出的，证明了希尔伯特定理。

你们看到，这是多么简单。但还要证明基本引理，而这一点，如同上一章一样，要经过一段漫长而艰难的路途。

§3 关于线性方程的引理

为了证明基本引理，我们要建立一系列辅助命题。

首先建立一些线性方程在整数范围内的解组个数的估计。或许，这节的引理自有一番独特的趣味，它们不属于我们要解决的问题。

引理一 设方程

$$a_1z_1+a_2z_2=m \tag{5}$$

中的数 a_1, a_2, m 是整数, $|a_1|\leqslant|a_2|\leqslant A$, 且 $(a_1, a_2)=1$。则方程 (5) 满足不等式 $|z_1|\leqslant A$, $|z_2|\leqslant A$ 的解数不超过 $3A/|a_2|$。

证明* 因 $|a_1|\leqslant|a_2|$, 且 $(a_1, a_2)=1$, 故 $a_2\neq 0$。不妨设 $a_2>0$, 否则, 只要在每组解中用 $-z_2$ 代替 z_2。

设 z_1, z_2 和 z_1', z_2' 是方程(5)的两组不同的解, 则由

$$a_1z_1+a_2z_2=m,$$

$$a_1z_1'+a_2z_2'=m,$$

得

$$a_1(z_1-z_1')=a_2(z_2'-z_2)。$$

这等式左边应被 a_2 整除, 但 $(a_1, a_2)=1$, 故 z_1-z_1' 应被 a_2 整除。$z_1\neq z_1'$, 说明 z_1-z_1' 是 a_2 的非 0 倍数, 故对(5)的任两组不同解 z_1, z_2 和 z_1', z_2', 必有

$$|z_1-z_1'|\geqslant a_2。$$

在方程(5)的每组解 z_1, z_2 中, 称 z_1 为第一员, z_2 为第二员。显然, 方程(5)满足条件 $|z_1|\leqslant A$, $|z_2|\leqslant A$ 的解数不大于落在段 $(-A, A)$ 中第一员的个数 t。上面已证得两个第一员相距不小于 a_2, 故其最大和最小的差不小于 $a_2(t-1)$, 另一方面, 因这差不超过 $2A$, 故

$$a_2(t-1)\leqslant 2A,$$

$$t-1\leqslant\frac{2A}{a_2},$$

$$t\leqslant\frac{2A}{a_2}+1\leqslant\frac{3A}{a_2}。$$

$\left(\text{因为, 根据假设 } a_2\leqslant A, \text{ 说明 } 1\leqslant\frac{A}{a_2}\right)$引理一证毕。

---

* 原书有错, 现已改正。——译者注

引理二 设方程

$$a_1z_1+a_2z_2+\cdots+a_lz_l=m \tag{6}$$

中的数 a_i 和 m 都是整数，

$$|a_i|\leqslant A \quad (1\leqslant i\leqslant l),\ (a_1,\ a_2,\ \cdots,\ a_l)=1。$$

则这方程满足不等式 $|z_i|\leqslant A(1\leqslant i\leqslant l)$ 的解数不超过

$$c(l)\frac{A^{l-1}}{H},$$

其中 H 是 $|a_1|,\ |a_2|,\cdots,|a_l|$ 中的最大者，而 $c(l)$ 是只与 l 有关的常数。

证明 显然，当 $l=2$ 时，引理二即引理一(取 $c(2)=3$)，故当 $l=2$，引理二已证。设 $l\geqslant 3$ 时，对 $l-1$，引理二成立。

因为编号的次序是无关紧要的，故可设 $|a_l|$ 是 $|a_1|,\ |a_2|,\ \cdots,\ |a_l|$ 中的最大者，即 $H=|a_l|$。

分两种情况讨论：

1) $a_1=a_2=\cdots=a_{l-1}=0$，因为 $(a_1,\ a_2,\ \cdots,\ a_l)=1$，故 $|a_l|=H=1$。则方程变成 $\pm z_l=m$。显然，在这方程中未知数 $z_1,\ z_2,\ \cdots,\ z_{l-1}$ 的每一个可取段 $(-A,\ A)$ 中的任意整数，即都有不多于 $2A+1\leqslant 3A$ 个选择，而 z_l 至多只取一个值，因此，方程(6)满足不等式 $|z_i|\leqslant A(1\leqslant i\leqslant l)$ 的解数不超过

$$(3A)^{l-1}=c(l)\cdot A^{l-1}=c(l)\cdot\frac{A^{l-1}}{H},$$

这时，引理二得证。

2) $a_1,\ a_2,\ \cdots,\ a_{l-1}$ 中至少有一个不是 0，令

$$(a_1,\ a_2,\ \cdots,\ a_{l-1})=\delta。$$

用 H' 表示数

$$\frac{|a_i|}{\delta} \quad (1\leqslant i\leqslant l-1)$$

中的最大者，并设 $z_1,\ z_2,\ \cdots,\ z_l$ 满足方程(6)和不等式 $|z_i|\leqslant A$

$(1\leqslant i\leqslant l)$。令

$$\frac{a_1}{\delta}z_1+\frac{a_2}{\delta}z_2+\cdots+\frac{a_{l-1}}{\delta}z_{l-1}=m'。\qquad(7)$$

由此得 $$a_1z_1+a_2z_2+\cdots+a_{l-1}z_{l-1}=\delta m'。$$

这时，显然

$$\delta m'+a_lz_l=m,\qquad(8)$$

而且 $$|\delta m'|\leqslant\sum_{i=1}^{l-1}|a_i||z_i|\leqslant\delta lH'A,$$

故得 $$|m'|\leqslant lH'A。$$

于是，如果数 $z_1, z_2, \cdots, z_l$ 满足方程(6)和不等式 $|z_i|\leqslant A$ $(1\leqslant i\leqslant l)$，则存在整数 m'，它满足方程式(7)和(8)，而且 $|m'|\leqslant lH'A$。但在方程(8)中，显然 $\delta\leqslant|a_l|$ 且 $(\delta, a_l)=1$（否则 $(a_1, a_2, \cdots, a_l)>1$）。故方程(8)的满足 $|m'|\leqslant lH'A$，$|z_l|\leqslant A<lH'A$ 的解数（未知数是 m', z_l），由引理一，不超过 $3lH'A/|a_l|$。对每个这样的 m'，由于引理二对 $l-1$ 个未知数成立，方程(7)的满足 $|z_i|\leqslant A$ 的整数解不多于 $c(l)\frac{A^{l-2}}{H'}$。

综上所述，方程(6)满足不等式 $|z_i|\leqslant A$，$1\leqslant i\leqslant l$ 的解数不超过

$$\frac{3lH'A}{|a_l|}\cdot c(l)\cdot\frac{A^{l-2}}{H'}=c(l)\cdot\frac{A^{l-1}}{|a_l|}=c(l)\cdot\frac{A^{l-1}}{H}。$$

引理二证毕*。

现在研究一切形为

$$a_1z_1+a_2z_2+\cdots+a_lz_l=0\qquad(9)$$

的方程组成的方程簇，其中 $|a_i|\leqslant A(1\leqslant i\leqslant l)$，而且所有 a_i 都是整数。设 B 是整数，它和 A 用不等式 $1\leqslant A\leqslant B\leqslant c(l)A^{l-1}$

---

* 你们也许注意到最后一个估计式中，$c(l)$ 在不同位置有不同的值。我在前面已事先告诉你们关于这个符号的这种用法。

联系，并设$l>2$。我们现在要估计这方程簇中解$z_i(|z_i|\leqslant B, 1\leqslant i\leqslant l)$个数的总和。

1° 首先考虑$a_1=a_2=\cdots=a_l=0$时的方程(9)(它在簇中)，并估计其满足不等式$|z_i|\leqslant B(1\leqslant i\leqslant l)$的解数。显然，任何数组$z_1, z_2, \cdots, z_l$，只要满足不等式$|z_1|\leqslant B, |z_2|\leqslant B, \cdots, |z_l|\leqslant B$，也必满足我们的方程。因$(-B, B)$中至多$2B+1$个数，则每个$z_i$至多可取$2B+1$个不同的数值，因而，我们感兴趣的数组$z_1, z_2, \cdots, z_l$的个数不超过

$$(2B+1)^l\leqslant(3B)^l=c(l)B^l。$$

又依假设，$B\leqslant c(l)A^{l-1}$，故$c(l)B^l=c(l)B^{l-1}\cdot B\leqslant c(l)(AB)^{l-1}$。因而，当$a_1=a_2=\cdots=a_l=0$时，方程(9)至多有$c(l)(AB)^{l-1}$个解。

2° 如果系数a_i中至少有一个非0，则存在最大公因数$(a_1, a_2, \cdots, a_l)=\delta$。首先设$\delta=1$，$H$是$|a_1|, |a_2|, \cdots, |a_l|$中的最大者。显然，$H$是段$(1, A)$中的一个整数，说明$H$或者在$A$和$\frac{A}{2}$之间，或者在$\frac{A}{2}$和$\frac{A}{4}$之间，或者在$\frac{A}{4}$和$\frac{A}{8}$之间等等，一般的，可找到整数$m\geqslant 0$，使得

$$\frac{A}{2^{m+1}}<H\leqslant\frac{A}{2^m}。\tag{10}$$

当$\delta=1$和H满足(10)时，由引理二，形为(9)的方程的解$z_i(|z_i|\leqslant B)$的个数不超过

$$c(l)\frac{B^{l-1}}{H}\leqslant c(l)\frac{B^{l-1}}{A/2^{m+1}}=\frac{c(l)B^{l-1}2^m}{A}。$$

另一方面，由不等式(10)得

$$|a_i|\leqslant\frac{A}{2^m}\quad(1\leqslant i\leqslant l),\tag{11}$$

说明满足不等式(10)的形为(9)的方程数不多于满足条件

(11)的同一类型的方程数，即至多

$$\left(2\times\frac{A}{2^m}+1\right)^l\leqslant\left(3\times\frac{A}{2^m}\right)^l=c(l)\cdot A^l\cdot 2^{-ml}。$$

因此，对 $\delta=1$ 且 $\frac{A}{2^{m+1}}<H\leqslant\frac{A}{2^m}$ 的形为(9)的方程，解 $|z_i|\leqslant B$ 的个数之和不超过

$$c(l)\frac{B^{l-1}2^m}{A}\cdot c(l)A^l\cdot 2^{-ml}=c(l)(AB)^{l-1}\cdot 2^{-(l-1)m}。$$

把这些估计式依 $m\geqslant 0$ 相加，则得以下结论：$|a_i|\leqslant A$ $(1\leqslant i\leqslant l)$，且 $\delta=1$ 的形为(9)的方程的解 $|z_i|\leqslant B$ 个数之和不超过

$$c(l)(AB)^{l-1}。$$

3° 最后计算当 $\delta>1$ 时方程的解数。这时，方程(9)显然等价于方程

$$\frac{a_1}{\delta}z_1+\frac{a_2}{\delta}z_2+\cdots+\frac{a_l}{\delta}z_l=0,$$

即又是形为(9)的方程，其中

$$\left(\frac{a_1}{\delta},\ \frac{a_2}{\delta},\ \cdots,\ \frac{a_l}{\delta}\right)=1,$$

而数 A 必须用 $\frac{A}{\delta}$ 代替。这样，对固定的 δ，所有这样的方程满足 $|z_i|\leqslant B$ 的解的个数总和，由 2°，不超过*

$$c(l)\left(\frac{A}{\delta}\cdot B\right)^{l-1}=c(l)\frac{(AB)^{l-1}}{\delta^{l-1}}。$$

依所有可能的值 $\delta(1\leqslant\delta\leqslant A)$ 把这些表达式加起来，则我们得到形为(9)的方程当 $|a_i|\leqslant A(1\leqslant i\leqslant l)$ 和 a_i 不全为 0 时，解的

---

* 因为现在的 A 用较小的数 $\frac{A}{\delta}$ 代替，前提条件 $B\leqslant c(l)A^{l-1}$ 可能不成立。但不难检验，在 2° 的论证中，我们没有用到这个前提条件，故 2° 的结果不依赖于这个条件。

个数之和不超过

$$c(l)(AB)^{l-1}\sum_{\delta=1}^{A}\frac{1}{\delta^{l-1}}<c(l)(AB)^{l-1}\cdot\frac{l-1}{l-2}{}^{*}$$
$$=c(l)(AB)^{l-1}。$$

把这和1°关于情况 $a_1=a_2=\cdots=a_l=0$ 的估计结果比较，归结为如下结论：

引理三. 设 $l>2$, $1\leqslant A\leqslant B\leqslant c(l)A^{l-1}$，则对于形为

$$a_1z_1+a_2z_2+\cdots+a_lz_l=0 \tag{9}$$

的所有方程，其中 $|a_i|\leqslant A(1\leqslant i\leqslant l)$，满足 $|z_i|\leqslant B(1\leqslant i\leqslant l)$ 的解的个数总和不超过

$$c(l)(AB)^{l-1}。$$

§4 另两个引理

证明基本引理之前，还需要建立两个特殊类型的辅助命

---

* 这里利用了不等式

$$\sum_{n=1}^{A}\frac{1}{n^{q+1}}<\frac{q+1}{q}$$

对任意自然数 q 和 A 成立（q 表示数 $l-2$，因 $l>2$，它是正的）。其简单证明是：当 $n\geqslant1$ 时

$$\frac{1}{n^q}-\frac{1}{(n+1)^q}=\frac{(n+1)^q-n^q}{n^q\cdot(n+1)^q}=\frac{n^q+qn^{q-1}+\cdots+1-n^q}{n^q\cdot(n+1)^q}$$
$$\geqslant\frac{qn^{q-1}}{n^q\cdot(n+1)^q}>\frac{q}{(n+1)^{q+1}},$$

故
$$\frac{1}{(n+1)^{q+1}}<\frac{1}{q}\left\{\frac{1}{n^q}-\frac{1}{(n+1)^q}\right\}。$$

在这些不等式中，依次令 $n=1, 2, \cdots, A-1$，并把所有这些不等式加起来，得到

$$\sum_{n=2}^{A}\frac{1}{n^{q+1}}<\frac{1}{q}\left(1-\frac{1}{A^q}\right)<\frac{1}{q},$$

故
$$\sum_{n=1}^{A}\frac{1}{n^{q+1}}<1+\frac{1}{q}=\frac{q+1}{q},$$

这就是所要证明的。

题。它们在形式关系和思想方法上都很简单，但要掌握它们，仍会使你们感到相当的困难，因为这里涉及到一切可能的组合数的估计，其结构是相当繁杂的。这种抽象组合的困难性在于：它很难使用数学符号，更多地要用语言叙述。当然，这是叙述的困难性，而不是问题本身的困难性。我将尽可能具体地跟你们讲清提出的问题以及它的解决办法。

用 A 表示有限数集，它所含的数可能有的相等；如果数 a 在集合 A 中出现 λ 次，我们称它的重数等于 λ。设 $a_1, a_2, \cdots, a_r$ 是 A 中不同的数，而 $\lambda_1, \lambda_2, \cdots, \lambda_r$ 是相应的重数（故集合 A 含有 $\sum_{i=1}^{r}\lambda_i$ 个数），设 B 是另一个同类型的集合，其不同的数是 $b_1, b_2, \cdots, b_s$，而重数分别是 $\mu_1, \mu_2, \cdots, \mu_s$。研究方程

$$x+y=c, \tag{12}$$

其中 c 是已知数，而 x 和 y 是未知数。对这个方程，我们感兴趣的解 x, y 要求 x 是取自集合 A 的数（我们将简单地写为 $x\in A$），而 y 是取自集合 B 的数（$y\in B$）。如果数 $x=a_i, y=b_k$ 满足方程(12)，则得到所求类型的 $\lambda_i\mu_k$ 个解，因为在 A 中的 λ_i 个"同样"的数 a_i 的任一个和 B 中的 μ_k 个"同样"的数 b_k 的任一个结合都是所求类型的解。但 $\lambda_i\mu_k\leqslant\frac{1}{2}(\lambda_i^2+\mu_k^2)$*，故方程(12)中 $x=a_i, y=b_k$ 这样的解数不超过 $\frac{1}{2}(\lambda_i^2+\mu_k^2)$。由此得到方程(12)的一切解 $x\in A, y\in B$ 的个数不超过和 $\sum\frac{1}{2}\cdot(\lambda_i^2+\mu_k^2)$，这里对一切使 $a_i+b_k=c$ 的符号 i, k 取和。如 λ_i^2 依

---

* "几何平均不超过算术平均"的最简单证明是

$$0\leqslant(\lambda_i-\mu_k)^2=\lambda_i^2+\mu_k^2-2\lambda_i\mu_k,$$

由此得

$$2\lambda_i\mu_k\leqslant\lambda_i^2+\mu_k^2.$$

所有 i 取和，而 μ_k^2 依所有 k 取和，则和增大(要知道，每个 b_k 至多只与一个 a_i 结合)。故最后得到方程(12)的解 $x\in A$, $y\in B$ 的个数不超过

$$\frac{1}{2}\left(\sum_{i=1}^{r}\lambda_i^2+\sum_{k=1}^{s}\mu_k^2\right)。$$

另一方面，研究方程

$$x-y=0, \tag{13}$$

求它的解 $x\in A$, $y\in A$ 的个数。显然，每个这样的解都有形式 $x=y=a_i(1\leqslant i\leqslant r)$；对给定的 i 有 λ_i^2 个解，因 x 和 y 可以独立地取在 A 中的 λ_i 个相同的数 a_i 的任一个。因此方程(13)的解 $x\in A$, $y\in A$ 的总个数等于 $\sum_{i=1}^{r}\lambda_i^2$；同样的，这方程的解 $x\in B$, $y\in B$ 的总个数等于 $\sum_{k=1}^{s}\mu_k^2$。综上所述，归结为以下的结论：

引理四 方程

$$x+y=c,\ x\in A,\ y\in B$$

的解数不超过方程

$$x-y=0,\ x\in A,\ y\in A$$

与

$$x-y=0,\ x\in B,\ y\in B$$

的解数之和的一半。

在集合 A 和 B 重合的特殊情况下，得

推论 方程

$$x+y=c,\ x\in A,\ y\in A$$

的解数不超过

$$x-y=0,\ x\in A,\ y\in A$$

的解数。

现在设 k 和 s 是任意的自然数，$k\cdot 2^s=l$，研究方程

$$x_1+x_2+\cdots+x_l=c。$$

再设 $A_1, A_2, \cdots, A_l$ 是有限数集，集合 A_i $(1\leqslant i\leqslant l)$ 含有重数分别是 $\lambda_{i_1}, \lambda_{i_2}, \cdots$ 的相互不同的数 $a_{i_1}, a_{i_2}, \cdots$，我们感兴趣的是方程

$$x_1+x_2+\cdots+x_l=c,\ x_i\in A_i\quad(1\leqslant i\leqslant l)\tag{14}$$

的解数。如果令

$$x_1+x_2+\cdots+x_{\frac{l}{2}}=x,\ x_{\frac{l}{2}+1}+\cdots+x_l=y$$

$\left(\frac{l}{2}\text{是整数}\right)$，则所给方程变成

$$x+y=c,$$

只要了解数 x 和 y 应属于怎样的集合，就可用引理四。因为 $x_i\in A_i(1\leqslant i\leqslant l)$，$x$ 是形为 $z_1+z_2+\cdots+z_{\frac{l}{2}}$ 的任何数，其中 $z_i\in A_i\left(1\leqslant i\leqslant\frac{l}{2}\right)$，而 y 是同样形式的数，但 $z_i\in A_{\frac{l}{2}+i}$ $\left(1\leqslant i\leqslant\frac{l}{2}\right)$。因此，由引理四，方程(14)的解数不超过方程

$$x-y=0\tag{15}$$

在下列两个假设之下，解的个数和的一半。

1)
$$x=z_1+z_2+\cdots+z_{\frac{l}{2}},$$
$$y=z'_1+z'_2+\cdots+z'_{\frac{l}{2}},$$

其中

$$z_i\in A_i,\quad z'_i\in A_i,\quad 1\leqslant i\leqslant\frac{l}{2}。\tag{16}$$

2) x 和 y 有同样的形式，但

$$z_i\in A_{\frac{l}{2}+i},\quad z'_i\in A_{\frac{l}{2}+i},\quad 1\leqslant i\leqslant\frac{l}{2}。\tag{17}$$

在这两种情况下，方程(15)都可写成

$$(z_1-z_1')+(z_2-z_2')+\cdots+(z_{\frac{l}{2}}-z'_{\frac{l}{2}})=0。\qquad (18)$$

故得结论：方程(14)的解数不超过方程(18)在假设(16)和(17)下解数和的一半，即不超过方程

$$\sum_{i=1}^{l/2}(z_i-z_i')=0,\quad z_i\in A_i,\quad z_i'\in A_i,\quad 1\leqslant i\leqslant\frac{l}{2}\qquad (18')$$

和 $$\sum_{i=1}^{l/2}(z_i-z_i')=0,\quad z_i\in A_{\frac{l}{2}+i},\quad z_i'\in A_{\frac{l}{2}+i},\quad 1\leqslant i\leqslant\frac{l}{2}\qquad (18'')$$

的解数和的一半。

方程(18)左边有 $\frac{l}{2}$ 个被加项，即比原先的方程(14)少一半。令

$$\sum_{i=1}^{l/4}(z_i-z_i')=x,\quad \sum_{i=\frac{l}{4}+1}^{l/2}(z_i-z_i')=y,$$

方程(18)化成形式

$$x+y=0。$$

还可以用引理四。完全如同从方程(14)到方程(18)那样，由(18)归结为方程

$$\sum_{i=1}^{l/4}(u_i+u_i'-u_i''-u_i''')=0。\qquad (19)$$

现在我们必须研究下列(已经有四种)假设下，这个方程的解数之和：

1) $u_i,\ u_i',\ u_i'',\ u_i'''\in A_i$,

2) $u_i,\ u_i',\ u_i'',\ u_i'''\in A_{\frac{l}{4}+i}$,

3) $u_i,\ u_i',\ u_i'',\ u_i'''\in A_{\frac{l}{2}+i}$, $\qquad\left(1\leqslant i\leqslant\frac{l}{4}\right)$。

4) $u_i,\ u_i',\ u_i'',\ u_i'''\in A_{\frac{3l}{4}+i}$

因为 $l=k\cdot 2^s$，故可重复这个过程 s 次，最后显然归结为方程

$$\sum_{i=1}^{k}\{y_i^{(1)}+y_i^{(2)}+\cdots+y_i^{(2^{s-1})}-y_i^{(2^{s-1}+1)}-\cdots-y_i^{(2^s)}\}=0, \quad (20)$$

而且必须研究在 2^s 种不同假设下，即

1) $y_1^{(j)}\in A_1,\ y_2^{(j)}\in A_2,\ \cdots,\ y_k^{(j)}\in A_k,$

2) $y_1^{(j)}\in A_{k+1},\ y_2^{(j)}\in A_{k+2},\ \cdots,\ y_k^{(j)}\in A_{2k},\quad (1\leqslant j\leqslant 2^s)$

…………………………………………

2^s) $y_1^{(j)}\in A_{k\cdot 2^s-k+1},\ \cdots,\ y_k^{(j)}\in A_{k\cdot 2^s},$

之时，方程的解数之和。

令 $\quad y^{(j)}=y_1^{(j)}+y_2^{(j)}+\cdots+y_k^{(j)}\quad (1\leqslant j\leqslant 2^s);$

则方程(20)归结为简单的形式

$$y^{(1)}+y^{(2)}+\cdots+y^{(2^{s-1})}-y^{(2^{s-1}+1)}-\cdots-y^{(2^s)}=0。\quad (21)$$

这里，涉及到方程(21)在下列用参数 $m(0\leqslant m\leqslant 2^s-1)$ 的 2^s 个不同的值区别的假设下，解数之和：

$$y^{(j)}=y_1^{(j)}+y_2^{(j)}+\cdots+y_k^{(j)},$$

其中 $\quad y_1^{(j)}\in A_{mk+1},\ y_2^{(j)}\in A_{mk+2},\ \cdots,\ y_k^{(j)}\in A_{(m+1)k}$

$$(j=1,\ 2,\ \cdots,\ 2^s)。$$

因而，我们可把我们的最后结论归纳为如下的命题：

引理五 方程

$$x_1+x_2+\cdots+x_l=c,\ x_i\in A_i,\quad 1\leqslant i\leqslant l=k\cdot 2^s \quad (14)$$

的解数不超过方程

$$y^{(1)}+y^{(2)}+\cdots+y^{(2^{s-1})}-y^{(2^{s-1}+1)}-\cdots-y^{(2^s)}=0,\quad (21)$$

$$y^{(j)}=y_1^{(j)}+y_2^{(j)}+\cdots+y_k^{(j)},$$

$$y_1^{(j)}\in A_{mk+1},\ y_2^{(j)}\in A_{mk+2},\ \cdots,\ y_k^{(j)}\in A_{(m+1)k}$$

$$(j=1,\ 2,\ \cdots,\ 2^s)$$

在 $m=0,\ 1,\ \cdots,\ 2^s-1$ 的前提下的解数之和。

显然，引理四是引理五当 $k=s=1,\ l=2$ 的特殊情况。

至此，准备工作完毕。下面，可以开始证明基本引理了。

§5 基本引理的证明

我们将用关于 n 的归纳法来证明基本引理。利用归纳证明，经常在加强被证明命题的同时，在本质上，用已知的方法使证明变得容易（实际上，有时是先使证明成为可实行的）。其原因不难理解。在归纳法证明中，假设命题当 $n-1$ 时成立，再来证明它当 n 时也成立；因此，命题越强，在 $n-1$ 的情况下，所给的条件也越多，而对数 n，要证明的东西也越多，但在许多问题中，条件较多显得更为重要。

在现在的情况下，我们感兴趣的问题是估计方程 $x_1^n+x_2^n+\cdots+x_k^n=\hat{m}(1\leqslant\hat{m}\leqslant N)$ 的解数（这时，就其问题本身的意义，$0\leqslant x_i\leqslant\hat{m}^{\frac{1}{n}}\leqslant N^{\frac{1}{n}}$）。但 x^n 是 n 次多项式

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n$$

的最简单的特殊情况。为了以后讨论的方便，我们用更一般的方程

$$f(x_1)+f(x_2)+\cdots+f(x_k)=\hat{m} \tag{22}$$

代替方程(1)，并假设未知数满足更宽的条件

$$|x_i|\leqslant N^{\frac{1}{n}}\quad(1\leqslant i\leqslant k)。$$

方程(22)的解数将比真正需要的更多，但这个推广一确立，从证明中立即可看出，能用归纳法！这时，对 $\hat{m}\leqslant N$，用 $r_k(\hat{m})$ 表示方程(22)满足条件 $|x_i|\leqslant N^{\frac{1}{n}}$，$1\leqslant i\leqslant k$ 的解数。在此，为了便于运用归纳法，当然允许多项式 $f(x)$ 的系数任意地附加条件（只要附加的条件包含 $f(x)=x^n$ 的情况）。我们要证明下列命题：

设多项式 $f(x)$ 满足

$$|a_i| \leqslant c(n) N^{\frac{i}{n}} \quad (0 \leqslant i \leqslant n), \tag{23}$$

则可选择 $k=k(n)$ 使

$$r_k(\hat{m}) < c(n) N^{\frac{k}{n}-1} \quad (1 \leqslant \hat{m} \leqslant N)。$$

因为不等式(23)当 $f(x)=x^n$，取 $c(n)=1$ 时，显然成立，故这命题实际上加强了基本引理。

首先考虑 $n=1$，$f(x)=a_0x+a_1$ 的情况，令 $k(1)=2$，则方程(22)归结为

$$a_0(x_1+x_2)=\hat{m}-2a_1,$$

而且，它的解满足 $|x_1| \leqslant N$，$|x_2| \leqslant N$。因此，x_1 至多只有 $2N+1 \leqslant 3N$ 个值，而每个 x_1 至多只对应一个 x_2，故

$$r_2(\hat{m}) \leqslant 3N,$$

即命题当 $n=1(k=2)$ 时已证得。

现在假设 $n>1$，命题对 $n-1$ 成立，令 $k(n-1)=k'$，并取

$$k=k(n)=2n\cdot 2^{[4\log_2 k']},$$

其中指数的符号表示不超过 $4\log_2 k'$ 的最大整数。以后，为简便起见，令 $[4\log_2 k']-1=s$，则

$$k=2n\cdot 2^{s+1}。\tag{24}$$

为估计方程(22)的解数 $r_k(\hat{m})$，首先利用引理四，令

$$x=\sum_{i=1}^{k/2} f(x_i), \quad y=\sum_{i=\frac{k}{2}+1}^{k} f(x_i)。$$

集合 A（这时，集合 B 和它重合）由形如

$$\sum_{i=1}^{k/2} f(x_i)$$

的一切数组成，其中 $|x_i| \leqslant N^{\frac{1}{n}}$，$1 \leqslant i \leqslant \frac{k}{2}$。由引理四的推论，$r_k(\hat{m})$ 不超过方程 $x-y=0$ 当 $x\in A$，$y\in A$，即

$$x=\sum_{i=1}^{k/2} f(x_i), \quad y=\sum_{i=1}^{k/2} f(y_i),$$

$$|x_i| \leqslant N^{\frac{1}{n}}, \ |y_i| \leqslant N^{\frac{1}{n}} \quad \left(1 \leqslant i \leqslant \frac{k}{2}\right)$$

时的解数。用另一句话说，$r_k(\hat{m})$不超过方程

$$\sum_{i=1}^{k/2}\{f(x_i)-f(y_i)\}=0 \tag{25}$$

当 $|x_i| \leqslant N^{\frac{1}{n}}$, $|y_i| \leqslant N^{\frac{1}{n}}\left(1 \leqslant i \leqslant \frac{k}{2}\right)$时的解数。令 $x_i-y_i=h_i\left(1 \leqslant i \leqslant \frac{k}{2}\right)$并用数组 y_i, h_i 代替未知数 x_i, y_i, 这时，允许 y_i 和 $h_i\left(1 \leqslant i \leqslant \frac{k}{2}\right)$是段$(-2N^{\frac{1}{n}}, 2N^{\frac{1}{n}})$中一切可能的整数。这样，所研究的方程的解数只会增加。那么，方程(25)的每个被加项$f(x_i)-f(y_i)$用表达式

$$\begin{aligned} f(y_i+h_i)-f(y_i) &= \sum_{v=0}^{n-1} a_v\{(y_i+h_i)^{n-v}-y_i^{n-v}\} \\ &= \sum_{v=0}^{n-1} a_v \sum_{t=1}^{n-v}\binom{n-v}{t} h_i^t y_i^{n-v-t} \end{aligned}$$

代替。为了改变求和的变数，令

$$v+t=u,$$

则
$$n-v-t=n-u, \ t=u-v。$$

故得

$$\begin{aligned} f(y_i+h_i)-f(y_i) &= h_i \sum_{v=0}^{n-1} a_v \sum_{u=v+1}^{n}\binom{n-v}{u-v} h_i^{u-v-1} y_i^{n-u} \\ &= h_i \sum_{u=1}^{n} y_i^{n-u} \sum_{v=0}^{u-1} a_v \binom{n-v}{u-v} h_i^{u-v-1} \\ &= h_i \sum_{u=1}^{n} a_{iu} y_i^{n-u} = h_i \phi_i(y_i), \end{aligned}$$

其中
$$\phi_i(y)=\sum_{u=1}^{n} a_{iu} y^{n-u}$$

是系数 $\quad a_{iu}=\sum_{v=0}^{u-1}a_v\binom{n-v}{u-v}h_i^{u-v-1}\quad\left(1\leqslant i\leqslant\frac{k}{2}\right)$

只与 h_i 有关的 $n-1$ 次多项式。

故在新变量 y_i, h_i 之下，方程(25)具有形式

$$h_1\varphi_1(y_1)+h_2\varphi_2(y_2)+\cdots+h_{\frac{k}{2}}\varphi_{\frac{k}{2}}(y_{\frac{k}{2}})=0。\qquad(26)$$

在这个方程中，数 h_i 和 y_i 可取段 $(-2N^{\frac{1}{n}},2N^{\frac{1}{n}})$ 中的任意整数，但要记住，($n-1$ 次)多项式 $\varphi(y)$ 的系数只与 h 有关。

这样一来，到这一步，我们证明了：要估计的数 $r_k(\hat{m})$，不超过当 h_i 取遍满足 $|h_i|\leqslant 2N^{\frac{1}{n}}\left(1\leqslant i\leqslant\frac{k}{2}\right)$ 的一切可能值时，形为(26)的方程的解 y_i, $|y_i|\leqslant 2N^{\frac{1}{n}}\left(1\leqslant i\leqslant\frac{k}{2}\right)$ 的个数之和。

§6 续 上 节

现在研究形如(26)的一个方程，即认为数 $h_i\left(1\leqslant i\leqslant\frac{k}{2}\right)$ 是固定的。应用引理五于所研究的方程，在这里，数 $h_i\varphi_i(y_i)$ 起了未知数 x_i 的作用，而数 $\frac{k}{2}=2n\cdot 2^s$ 起了 l 的作用。为简便起见，令 $2n=k_0$。还要记住，数 h_i 不仅明显地出现在方程(26)中，而且隐含在多项式 $\varphi_i(y)$ 的系数中。数 $x_i=h_i\varphi_i(y_i)$ 所属的集合 A_i 由形如 $h_i\varphi_i(y_i)$ 的一切数组成，其中 h_i 有固定的常数值，而数 y_i 跑遍段 $(-2N^{\frac{1}{n}},2N^{\frac{1}{n}})$ 中的整数。

由引理五，满足刚刚描述的条件的方程(26)的解数不超过方程

$$y^{(1)}+y^{(2)}+\cdots+y^{(2^{s-1})}-y^{(2^{s-1}+1)}-\cdots-y^{(2^s)}=0\qquad(21)$$

在下列 2^s 个假设

$$\left.\begin{array}{l} y^{(j)}=y_1^{(j)}+y_2^{(j)}+\cdots+y_{k_0}^{(j)}, \\ y_i^{(j)}\in A_{mk_0+i} \quad (1\leqslant i\leqslant k_0) \end{array}\right\}(1\leqslant j\leqslant 2^s)$$

下的解数之和，其中，参数 $m=0, 1, \cdots, 2^s-1$，而 $A_r(1\leqslant r\leqslant 2^s)$ 是当 h_r 固定，y_r 满足 $|y_r|\leqslant 2N^{\frac{1}{n}}$ 时的形为 $h_r\varphi_r(y_r)$ 的数集。

对于 $m=0$ 的情况（只是选做为一例），在展开的形式中，方程(21)为

$$\begin{aligned} &\{y_1^{(1)}+y_2^{(1)}+\cdots+y_{k_0}^{(1)}\}+\{y_1^{(2)}+y_2^{(2)}+\cdots+y_{k_0}^{(2)}\} \\ &+\cdots+\{y_1^{(2^{s-1})}+y_2^{(2^{s-1})}+\cdots+y_{k_0}^{(2^{s-1})}\} \\ &-\{y_1^{(2^{s-1}+1)}+y_2^{(2^{s-1}+1)}+\cdots+y_{k_0}^{(2^{s-1}+1)}\} \\ &-\cdots-\{y_1^{(2^s)}+y_2^{(2^s)}+\cdots+y_{k_0}^{(2^s)}\}=0, \end{aligned}$$

或者，改变被加项的次序，得

$$\begin{aligned} &\{y_1^{(1)}+y_1^{(2)}+\cdots+y_1^{(2^{s-1})}-y_1^{(2^{s-1}+1)}-\cdots-y_1^{(2^s)}\} \\ &+\{y_2^{(1)}+y_2^{(2)}+\cdots+y_2^{(2^{s-1})}-y_2^{(2^{s-1}+1)}-\cdots-y_2^{(2^s)}\} \\ &+\cdots+\{y_{k_0}^{(1)}+y_{k_0}^{(2)}+\cdots+y_{k_0}^{(2^{s-1})}-y_{k_0}^{(2^{s-1}+1)}-\cdots-y_{k_0}^{(2^s)}\}=0; \end{aligned}$$

在这里，每个数 $y_i^{(j)}$ 都有 $h_i\varphi_i(v_i^{(j)})$ 的形式，其中 $|v_i^{(j)}|\leqslant 2N^{\frac{1}{n}}$。因此，最后的方程可写成

$$\begin{aligned} &h_1\{\varphi_1(v_1^{(1)})+\varphi_1(v_1^{(2)})+\cdots+\varphi_1(v_1^{(2^{s-1})}) \\ &\quad-\varphi_1(v_1^{(2^{s-1}+1)})-\cdots-\varphi_1(v_1^{(2^s)})\} \\ &+h_2\{\varphi_2(v_2^{(1)})+\cdots-\varphi_2(v_2^{(2^s)})\}+\cdots \\ &+h_{k_0}\{\varphi_{k_0}(v_{k_0}^{(1)})+\cdots-\varphi_{k_0}(v_{k_0}^{(2^s)})\}=0。 \end{aligned}$$

为简便起见，设

$$\begin{aligned} &\varphi_i(v_i^{(1)})+\varphi_i(v_i^{(2)})+\cdots+\varphi_i(v_i^{(2^{s-1})}) \\ &\quad-\varphi_i(v_i^{(2^{s-1}+1)})-\cdots-\varphi_i(v_i^{(2^s)})=z_i \\ &\qquad (1\leqslant i\leqslant k_0), \end{aligned}$$

则可把这方程变得更简单：

$$h_1z_1+h_2z_2+\cdots+h_{k_0}z_{k_0}=0。\tag{27}$$

这样的方程总共有 2^s 个，其全体可简写成形式

$$\sum_{i=1}^{k_0} h_{mk_0+i} z_{mk_0+i}=0 \quad (0\leqslant m\leqslant 2^s-1)。$$

但我们只限于研究方程(27)，它可视为这类方程的典型代表。为了估计我们感兴趣的解数，首先必须说明量 $\varphi_i(v_i^{(j)})$ 可在怎样的范围内取值。为此，记住(见 44 页)

$$\varphi_i(y)=\sum_{u=1}^{n} a_{iu} y^{n-u},$$

其中
$$a_{iu}=\sum_{v=0}^{u-1} a_v \binom{n-v}{u-v} h_i^{u-v-1} \quad \left(1\leqslant i\leqslant \frac{k}{2}\right)。$$

因此，利用以上的前提条件 $|a_v|\leqslant c(n)N^{\frac{v}{n}}$ 和 $|h_i|\leqslant 2N^{\frac{1}{n}}$ 得

$$|a_{iu}|\leqslant \sum_{v=0}^{u-1} c(n)N^{\frac{v}{n}}\binom{n-v}{u-v} c(n) N^{\frac{u-v-1}{n}}$$
$$=c(n)N^{\frac{u-1}{n}}\sum_{v=0}^{u-1}\binom{n-v}{u-v},$$

因 $u\leqslant n$，故

$$|a_{iu}|\leqslant c(n)N^{\frac{u-1}{n}}。 \tag{28}$$

但是，另一方面，因为 $|v_i^{(j)}|\leqslant 2N^{\frac{1}{n}}$，故 $|v_i^{(j)}|^{n-u}\leqslant c(n)N^{\frac{n-u}{n}}$，因此，

$$|a_{iu}|\cdot|v_i^{(j)}|^{n-u}\leqslant c(n)N^{\frac{u-1}{n}}N^{\frac{n-u}{n}}=c(n)N^{\frac{n-1}{n}}。$$

把这个界(具有另一个 $c(n)$)代入所有的 $\varphi_i(v_i^{(j)})$，因为多项式的项数为 n，故

$$|\varphi_i(v_i^{(j)})|\leqslant c(n)N^{\frac{n-1}{n}}, \quad 1\leqslant i\leqslant k_0\cdot 2^s, \quad 1\leqslant j\leqslant 2^s。$$

但每个 z_i 是 $2^s=c(n)$ 个形如 $\pm\varphi_i(v_i^{(j)})$ 的被加项的和，故

$$|z_i|\leqslant c(n)N^{\frac{n-1}{n}} \quad (1\leqslant i\leqslant 2^s)$$

(当然，具有另一个 $c(n)$)。故在方程(27)中，每个 z_i 只能取段

$(-c(n)N^{\frac{n-1}{n}}, c(n)N^{\frac{n-1}{n}})$中的整数。设$\overline{m}$是其中的一个数，一般地说，等式$z_i=\overline{m}$不只是以一种方式，而是以若干种方式存在，因为数$z_i$的定义是：同一个$z_i$的值可用$v_i^{(j)}(1\leqslant j\leqslant 2^s)$的不同选择得到(见第 46 页)。我们现在必须估计关系式$z_i=\overline{m}$的解数，即方程

$$\varphi_i(v_i^{(1)})+\cdots+\varphi_i(v_i^{(2^{s-1})})$$
$$-\varphi_i(v_i^{(2^{s-1}+1)})-\cdots-\varphi_i(v_i^{(2^s)})=\overline{m} \qquad (29)$$

的解数。这可用归纳法。

首先方程(29)化为

$$\varphi_i(v_i^{(1)})+\varphi_i(v_i^{(2)})+\cdots+\varphi_i(v_i^{(k')})$$
$$=\overline{m}-\varphi_i(v_i^{(k'+1)})-\cdots+\varphi_i(v_i^{(2^{s-1}+1)})+\cdots+\varphi_i(v_i^{(2^s)})。$$

因为$k'=k(n-1)>1$(我们已知$k(1)=2$)，故有

$$2^{s-1}=2^{[4\log_2 k']-2}>k'\,^{*},$$

从而，这样写是可以的。

把最后一个方程的右边记为m'，得

$$\varphi_i(v_i^{(1)})+\cdots+\varphi_i(v_i^{(k')})=m'。 \qquad (30)$$

关于数 $\qquad v_i^{(j)} \quad (k'+1\leqslant j\leqslant 2^s),$

选择一个确定的值(当然在段$[-2N^{\frac{1}{n}}, 2N^{\frac{1}{n}}]$中)，则$m'$也是确定的值。利用归纳法假设，因为所有必要的前提条件都成立，实际上，我们有

$$\varphi_i(y)=\sum_{u=1}^{n}a_{iu}y^{n-u},$$

是$n-1$次多项式，由(28)，

$$|a_{iu}|\leqslant c(n)N^{\frac{u-1}{n}}=c(n)\,(N^{\frac{n-1}{n}})^{\frac{u-1}{n-1}}, \qquad (31)$$

---

* 详细地说：$k'\geqslant 2$，$\log_2 k'\geqslant 1$，$3\log_2 k'\geqslant 3$，$[4\log_2 k']-2>4\log_2 k'-3\geqslant\log_2 k'$，故$2^{s-1}=2^{[4\log_2 k']-2}>k'$。

且不难得出 $$|m'| \leqslant c(n) N^{\frac{n-1}{n}}$$
（因 $\overline{m}$ 和每个 $\varphi_i(y_i^{(j)})$ 都满足这个不等式）。

在此，最后的不等式中，数 $c(n)N^{\frac{n-1}{n}}$ 起了数 N 的作用，而条件(31)是关于 $\varphi_i(y)$ 的系数的约束条件，即条件(23)中用 $n-1$ 代替 n。因此，一切前提条件成立，说明方程(30)对于 $|v_i^{(j)}| \leqslant 2N^{\frac{1}{n}} = 2(N^{\frac{n-1}{n}})^{\frac{1}{n-1}}$ 的解数不超过

$$c(n)(N^{\frac{n-1}{n}})^{\frac{k'}{n-1}-1} = c(n) N^{\frac{k'-n+1}{n}}。 \tag{32}$$

这个估计是对固定值 $v_i^{(k'+1)}, \cdots, v_i^{(2^s)}$ 得到的。显然，这样数组的总数不多于

$$(2N^{\frac{1}{n}}+1)^{2^s-k'} < c(n) N^{\frac{2^s-k'}{n}}。 \tag{33}$$

故方程(29)所要求类型的解数不超过(32)和(33)右边部分的积，即不超过

$$c(n) N^{\frac{2^s-n+1}{n}}。 \tag{34}$$

现在回到方程(27)，上面已经证明，每个 z_i 只能取段 $(-c(n)N^{\frac{n-1}{n}},\ c(n)N^{\frac{n-1}{n}})$ 中的一个值，现在又证明了这种值的"重数"(即它存在的前提下，选择 $y_i^{(j)}$ 的方法数）不超过(34)。

这个结论允许我们把所有问题归结为计算线性方程解的个数。实际上，在§5末，已把估计 $r_k(\hat{m})$ 归结为估计形为(26)的方程的解数；但方程(26)的解当 $|y_i| \leqslant 2N^{\frac{1}{n}}$ 时的个数，正如利用引理五所证明的，不超过 2^s 个形为(27)的已是线性的方程的解数，在此，未知数 z_i 的界也已得到。一些新的困难性（为转移到线性方程必须付出的代价）在于：新的未知数 z_i 应考虑为具有一定的重数（其上界也已得到）。

最后，我们不应忘记，所有的演算都有一个前提：数 h_i 是

选定的。因此，我们还应把这结果乘上所有可能的选择数。

那么，这一节的最后结论是：我们所估计的数 $r_k(\hat{m})$ 不超过方程

$$\sum_{i=1}^{k_0} h_{mk_0+i} z_{mk_0+i} = 0, \qquad (35)$$

当 z_i 是整数，$|z_i| \leqslant c(n) N^{\frac{n-1}{n}}$，重数 $\lambda_i \leqslant c(n) N^{\frac{2^s-n+1}{n}}$，$m$ 取遍 $0, 1, \cdots, 2^s-1$，而数 $h_r (1 \leqslant r \leqslant 2^s k_0)$ 相互独立地取遍段 $(-2N^{\frac{1}{n}}, 2N^{\frac{1}{n}})$ 中的整数时，解数的总和。

数 $r_k(\hat{m})$ 的这么一个估计，它的形式不涉及 $f(x)$，因而具有十分普遍的特征。

§7 结　论

在把问题归结为与特殊形式的多项式 $f(x)$ 无关的线性方程的解数估计之后，我们很容易借助于引理三得到结论。

用 Γ 表示数 $h_i \left(|h_i| \leqslant 2N^{\frac{1}{n}}, 1 \leqslant i \leqslant \frac{k}{2} \right)$ 的某一个确定组合，并用 $U_m(\Gamma)$ 表示在某个给定的 m 和这个组合之下，解 z_i 满足不等式 $|z_i| \leqslant c(n) N^{\frac{n-1}{n}}$，重数 $\lambda_i \leqslant c(n) N^{\frac{2^s-n+1}{n}}$ 时，方程(35)的解数。则由上节得到的结论

$$r_k(\hat{m}) \leqslant \sum_{\Gamma} \left\{ \sum_{m=0}^{2^s-1} U_m(\Gamma) \right\},$$

其中 Γ 取遍数 h_i 的一切容许组合。这可改写为

$$r_k(\hat{m}) \leqslant \sum_{m=0}^{2^s-1} \left\{ \sum_{\Gamma} U_m(\Gamma) \right\}。$$

但显然，立即可得出，对不同的 m，和 $\sum_{\Gamma} U_m(\Gamma)$ 没有什么不同（因对不同的 m，方程(35)没有什么不同），故又可写为

$$r_k(\hat{m}) \leqslant 2^s \sum_{\Gamma} U_0(\Gamma) = c(n) \sum_{\Gamma} U_0(\Gamma)。$$

这里，$U_0(\Gamma)$是方程

$$h_1 z_1 + h_2 z_2 + \cdots + h_{k_0} z_{k_0} = 0 \tag{36}$$

在数$h_i\left(|h_i| \leqslant 2N^{\frac{1}{n}},\ 1 \leqslant i \leqslant \frac{k}{2}\right)$有给定的组合$\Gamma$，$|z_i| \leqslant c(n) \cdot N^{\frac{n-1}{n}}$，且$z_i$具有重数$\lambda_i \leqslant c(n) N^{\frac{2^s-n+1}{n}}$时的解数。如果用$U_0^*(\Gamma)$表示同一方程在$z_i$都是一重的假设下的解数，则

$$U_0(\Gamma) \leqslant \{c(n) N^{\frac{2^s-n+1}{n}}\}^{k_0} U_0^*(\Gamma),$$

或者，回想到$k_0 = 2n$，得

$$U_0(\Gamma) \leqslant c(n) \cdot N^{2(2^s-n+1)} U_0^*(\Gamma),$$

故

$$r_k(\hat{m}) \leqslant c(n) \cdot N^{2(2^s-n+1)} \sum_{\Gamma} U_0^*(\Gamma)。 \tag{37}$$

我们接下去要指出，每个Γ是所有值$h_i\left(1 \leqslant i \leqslant \frac{k}{2}\right)$的某一个容许组合，同时，数$U_0^*(\Gamma)$完全由这些数$(1 \leqslant i \leqslant 2n)$的前$k_0 = 2n$个值确定，因为在方程(36)中只有这些数出现。选择某个确定的组合Γ后，我们也同样地确定了具有值$h_1, h_2, \cdots, h_{2n}$的某一个组合Γ'；但是，反过来，数$h_1, h_2, \cdots, h_{2n}$的某个确定组合Γ'选定后，它不是对应一个组合Γ，其它的数$h_i\left(2n < i \leqslant \frac{k}{2}\right)$可以"补选"出的方法数有多少，它对应的组合数也有多少。因为每个h_i应在段$(-2N^{\frac{1}{n}}, 2N^{\frac{1}{n}})$中，故一个组合$\Gamma'$对应的组合$\Gamma$不多于

$$c(n)\,(N^{\frac{1}{n}})^{\frac{k}{2}-2n} = c(n)\, N^{\frac{k}{2n}-2}$$

个。因而

$$\sum_{\Gamma} U_0^*(\Gamma) \leqslant c(n) N^{\frac{k}{2n}-2} \sum_{\Gamma'} U_0^*(\Gamma'),$$

其中 $U_0^*(\Gamma')$ 是方程(36)在满足条件: $|z_i|\leqslant c(n)N^{\frac{n-1}{n}}$, $1\leqslant i\leqslant 2n$, 数 h_i, $|h_i|\leqslant 2N^{\frac{1}{n}}$, $1\leqslant i\leqslant 2n$, 具有给定的组合 Γ' 时整数解 z_i 的个数。依所有这些组合, 把上式加起来, 由(37)得到(记住, $k=2n\cdot 2^{s+1}$)

$$r_k(\hat{m})\leqslant c(n)N^{2(2^s-n+1)}N^{\frac{k}{2n}-2}\sum_{\Gamma'}U_0^*(\Gamma')$$
$$=c(n)N^{2(2^{s+1}-n)}\sum_{\Gamma'}U_0^*(\Gamma')。\qquad (38)$$

最后, 利用引理三直接估计 $\sum_{\Gamma'}U_0^*(\Gamma')$, 令 $l=2n$, $A=2N^{\frac{1}{n}}$, $B=c(n)N^{\frac{n-1}{n}}$。容易验证, 引理三的前提条件都成立, 故得:

$$\sum_{\Gamma'}U_0^*(\Gamma')\leqslant c(n)(AB)^{2n-1}=c(n)N^{2n-1}。$$

因此, (38)变成

$$r_k(\hat{m})\leqslant c(n)N^{2(2^{s+1}-n)}\cdot N^{2n-1}$$
$$=c(n)N^{2(2^{s+1}-1)}=c(n)\cdot N^{\frac{k}{n}-1}。$$

终于完成了基本引理的证明, 从而, 证明了希尔伯特定理。

这个证明, 就其初等性而言, 多么美妙! 但是无疑的, 你们会感到它多么复杂! 因此, 值得花两、三个星期的时间, 亲自动手演算一番, 才能完全理解和掌握。数学家正是在克服诸如此类的困难之中, 成长和发展起来的。

［Ｇｅｎｅｒａｌ Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝数论的三颗明珠
作者＝［苏联］Ａ．Я．辛钦
页数＝５２
ＳＳ号＝４１０２５４１８５